



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 180386

**TO: Jeanine Goldberg**  
**Art Unit: 1634**  
**Location: REM-2D15&2C70**  
**Serial Number: 10/699941**

**Friday, March 10, 2006**

**From: Beverly Shears**  
**Location: Biotech-Chem Library**  
**REM 1A54**  
**Phone: 571-272-2528**  
**beverly.shears@uspto.gov**

### Search Notes

Your queries have completed processing. You may access an electronic version via eDAN (SCORE) and /or <http://es/ScoreAccessWeb>.

#### Protein Sequence Searches – February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (uniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

#### Published Applications Database - November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New).  
Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).



**THIS PAGE BLANK (USPTO)**

STIC-Biotech/ChemLib

180386

From: Goldberg, Jeanine  
Sent: Thursday, February 23, 2006 12:54 PM  
To: STIC-Biotech/ChemLib  
Subject: 10/699941- ataxia-

Hello-  
Please search SEQ ID NO: 3.

THANKS  
Jeanine

Jeanine Anne Goldberg  
1634  
571-272-0743  
REM 2D15  
Mailbox: 2C70

CRFES

RECEIVED  
FEB 23 2006  
STIC

3-2918NA

3/07  
STB

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: \_\_\_\_\_  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

Date completed: \_\_\_\_\_  
Searcher: Beverly e 2528  
Terminal time: \_\_\_\_\_  
Elapsed time: \_\_\_\_\_  
CPU time: \_\_\_\_\_  
Total time: \_\_\_\_\_  
Number of Searches: \_\_\_\_\_  
Number of Databases: \_\_\_\_\_

Search Site

\_\_\_\_\_ STIC  
\_\_\_\_\_ CM-1  
\_\_\_\_\_ Pre-S

Type of Search

\_\_\_\_\_ N.A. Sequence  
\_\_\_\_\_ A.A. Sequence  
\_\_\_\_\_ Structure  
\_\_\_\_\_ Bibliographic

Vendors

\_\_\_\_\_ IG  
\_\_\_\_\_ STN  
\_\_\_\_\_ Dialog  
\_\_\_\_\_ APS  
\_\_\_\_\_ Geninfo  
\_\_\_\_\_ SDC  
\_\_\_\_\_ DARC/Questel  
\_\_\_\_\_ Other CGN

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.7

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 03:15:24 ; Search time 14445 Seconds  
(without alignments)  
11482.810 Million cell updates/sec

Title: US-10-699-941-3  
Perfect score: 2918  
Sequence: 1 gcgcgcctctgcgcgcct.....ttctcaaaaaaaaaaaaaa 2918

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5881141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*

1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_sta.\*  
11: gb\_sy.\*  
12: gb\_un.\*  
13: gb\_vl.\*  
14: gb\_hfg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2918	100.0	2918	8	BC026217 Homo sapi
2	2896.6	99.3	4976	6	BD171164 Novel gen
3	2896.6	99.3	4976	6	BD183464 Novel gen
4	2896.6	99.3	4976	8	AB058775 Homo sapi
5	2887.6	99.0	3076	6	AR560886 Sequence
6	2870.2	98.4	2894	6	BD158743 Primer fo
7	2870.2	98.4	2894	6	AX881045 Sequence
8	2870.2	98.4	2894	6	AK027889 Sequence
9	2598.2	89.0	2759	6	AX747472 Sequence
10	2598.2	89.0	2759	8	AK092309 Homo sapi
11	2463	84.4	2787	8	AB052149 Macaca fa
12	2357.6	80.8	2786	6	CQ842483 Sequence
13	2357.6	80.8	2786	8	AK125457 Homo sapi
14	1815.8	62.2	2276	8	BC008736 Homo sapi
15	1438.2	49.3	134308	8	AC011488 Homo sapi
16	1438.2	49.3	210617	14	AC034201 Homo sapi
17	1325.8	45.4	2625	6	AX747461 Sequence
18	1325.8	45.4	2625	8	AK092298 Homo sapi

19	1114.4	38.2	1116	8	AY220297
20	1113	38.1	1113	6	AR560887 Sequence
21	923.2	31.6	997	6	CQ721138 Sequence
22	857.8	29.4	1629	9	AY349150 Mus muscu
23	857.8	29.4	3677	9	BC048903 Mus muscu
24	745.2	25.5	790	6	BD150580 Primer fo
25	745.2	25.5	790	6	AX870518 Sequence
26	579	19.8	579	6	CQ832574 Sequence
27	522.2	17.9	1821	8	AB046623 Macaca fa
28	470	16.1	560	6	BD155501 Primer fo
29	470	16.1	560	6	AX875439 Sequence
30	346.4	11.9	4507	9	BC094224 Mus muscu
31	341	11.7	4164	9	BC095978 Mus muscu
32	332	11.4	2242	9	BC091368 Rattus no
33	330.4	11.3	5656	8	AB002365 Human mRN
34	327.6	11.2	11690	8	AB050197 Homo sapi
35	326.6	11.2	4284	9	AK172939 Mus muscu
36	318.2	10.9	5654	6	AX330135 Sequence
37	318.2	10.9	5654	6	AX336310 Sequence
38	318.2	10.9	5654	6	AX336690 Sequence
39	315.8	10.8	2347	6	BD083730 Nucleic a
40	315.8	10.8	2347	6	BD097377 Nucleic a
41	311.4	10.7	2491	8	AY439213 Homo sapi
42	311	10.7	1804	5	AY423003 Danio rer
43	309.8	10.6	322	6	BD118568 EST and e
44	309.8	10.6	322	6	AR423015 Sequence
45	309.8	10.6	322	6	AX983709 Sequence

## ALIGNMENTS

## RESULT 1

## BC026217

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

BC026217 2918 bp mRNA linear PRI 29-JUN-2004  
Homo sapiens ataxia, cerebellar, Cayman type (caytaxin), mRNA (CDNA  
clone MGC:21400 IMAGE:4153341), complete cds.

BC026217 GI:20070729

MGC.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 2918)

Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.P., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Heiton, E., Kettelman, M., Madan, A., Rodrigues, S.,  
Sanchez, A., Whitting, M., Madan, A.C., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butcherfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,  
Schnerker, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences

proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2918)

Straussberg, R.

Direct Submission

Submitted (02-APR-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Goldberg, J.  
10/6/99 941  
Seg. ID 3

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgs.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabsa-r@mail.nih.gov](mailto:cgabsa-r@mail.nih.gov)  
Tissue Procurement: David N. Louis, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-ahgc.stanford.edu>  
Contact: (Dickson, Mark) [mdc@paxil.stanford.edu](mailto:mdc@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 20 Row: b Column: 2  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

## FEATURES

source

1..2918  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:21400 IMAGE:4153341"  
/tissue types="Brain, anaplastic oligodendroglioma with 1p/19q loss"  
/clone\_lib="NCI CGAP\_Brn67"  
/lab\_host="DH10B"  
/notes="Vector: pCMV-SPORT6"

gene

1..2918  
/genes="ATCAY"  
/note="synonyms: KIAA1872, BNIP-H, CLAC"  
/db\_xref="GeneID:85300"  
/db\_xref="MIM:608179"  
373..1488  
/genes="ATCAY"  
/codon\_start=1  
/product="caytaxin"

CDS

/protein\_id="AAH26217.1"  
/db\_xref="GI:20070730"  
/db\_xref="GeneID:85300"  
/db\_xref="MIM:608179"  
translation="MGTRATLNMENVKEDLDLRLPLPEETGVLLGSPVEDT  
SSPNFLNFGARXKTLVAPENISLQSGSLLSDFLPTDDLDINVDIETPD  
ETDLSFLNGNELEWDDTPVATANKMPGDSADLFGDTTDDGSAANRLMRTVIG  
EQHRDLHMIRPYMKVTHGGYEGELNAIVFAACFLPDSLSLPDYHIMENLFLYV  
ISLELLVAEDYIMVYLANGATPRRRPGIGWLKQMDRLRLKSLIIVHPSWF  
IRTVLAISRPFISVKINKIQYVHSLDLEQLIPMEHVQIPDCVLQYBEERLKARRS  
ARQPEFVLPSEKEVAPVENSALVSDQETSMS"

## ORIGIN

Query Match 100.0%; Score 2918; DB 8; Length 2918;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2918; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GCGAGCCTCTGCCAGCCCTGAGCTGGGAGAGACAGCTACCTCGGAGGCGAGGCGCGCA	60
Db	1		
Qy	61	GCGGCGGCGGATGAGAGGGGGCGCAGCCGCGCGCTGGGAGCCACCGCTAAC	120
Db	61		
Qy	121	CTTGCAACCCACCCCTCTGCACAAAGAGCTGGGGGGCTGGGCAACGTCGCCCTGGG	180
Db	121		
Qy	181	TGACCTTCCTCGGATCGAGATCGCCCTCGGAGCATCTCTCTCTCTCTCTCTCTCTGAA	240
Db	181		



Matches 2911; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

QY	1	GCCGAGCCTCTGCCAGCCTGAGCTGGGAAGACAGCTACCTCGGAGGACAGGGCGCGCA	60
Db	5	GCCGAGCCTCTGCCAGCCTGAGCTGGGAAGACAGCTACCTCGGAGGACAGGGCGCGCA	64
QY	61	GGCGGGGGGGATGAGAGGGGGGGAGCCGCGAGCCCGCGCTGGGGAGCCACCGCTAAC	120
Db	65	GGCGGGGGGGATGAGAGGGGGGGAGCCGCGAGCCCGCGCTGGGGAGCCACCGCTAAC	124
QY	121	CCTGCACCCACCCCTCTGACAAAGAGCTGGGGGGCTGGCCAGCTGGCCCTGGG	180
Db	125	CCTGCACCCACCCCTCTGACAAAGAGCTGGGGGGCTGGCCAGCTGGCCCTGGG	184
QY	181	TGACCTTCTCTGGATGAGATCCGCCCTCTGGAGCATCTCTTCTCTAGGCTCTGAA	240
Db	185	TGACCTTCTCTGGATGAGATCCGCCCTCTGGAGCATCTCTTCTCTAGGCTCTGAA	244
QY	241	GGCCCGGGAGCGTAGCGATGCCAGCTGCACCCGGGGCAGGGCTCGCCTTTGTTGCCA	300
Db	245	GGCCCGGGAGCGTAGCGATGCCAGCTGCACCCGGGGCAGGGCTCGCCTTTGTTGCCA	304
QY	301	GTAAAGGAGAGAGCTGTCTCAGCTGCAGAGGGTCTATCCCTGCTTCAAGCCAGTGCCTC	360
Db	305	GTAAAGGAGAGAGCTGTCTCAGCTGCAGAGGGTCTATCCCTGCTTCAAGCCAGTGCCTC	364
QY	361	TTCCAGCTCCCATGGGGACCAACGGAAGCAGCTCCGGATGGAACCGTGGACGTGAAG	420
Db	365	TTCCAGCTCCCATGGGGACCAACGGAAGCAGCTCCGGATGGAACCGTGGACGTGAAG	424
QY	421	GAGGAATGGCAGGACGAAGATTTTCCAGGCCATCTCCAGAGAGAGAGCGGGGTGGAATG	480
Db	425	GAGGAATGGCAGGACGAAGATTTTCCAGGCCATCTCCAGAGAGAGAGCGGGGTGGAATG	484
QY	481	CTTGGCAGCCCGTGAAGACATCTCTCTCCCAACACGCTAAATTTCAACGGAGCG	540
Db	485	CTTGGCAGCCCGTGAAGACATCTCTCTCCCAACACGCTAAATTTCAACGGAGCG	544
QY	541	CATCGTAAGAGGAAGCGCTGTGGCCACGAGATCAACATTTCTCGATCAGAGTGAG	600
Db	545	CATCGTAAGAGGAAGCGCTGTGGCCACGAGATCAACATTTCTCGATCAGAGTGAG	604
QY	601	GGTCCCTGTGTCCGATGATCTTTGGATACCCCTGATGACTGGATATTAACGTGGAT	660
Db	605	GGTCCCTGTGTCCGATGATCTTTGGATACCCCTGATGACTGGATATTAACGTGGAT	664
QY	661	GACATCGAGACCCCGATGAGACCGACTCGCTGGAGTTCTGGGGAATGGCAACGACTG	720
Db	665	GACATCGAGACCCCGATGAGACCGACTCGCTGGAGTTCTGGGGAATGGCAACGACTG	724
QY	721	GAGTGGGAAGACGACACCCCGTGGCCACCGCAAGAACATGCCGGGACAGCGCGGAT	780
Db	725	GAGTGGGAAGACGACACCCCGTGGCCACCGCAAGAACATGCCGGGACAGCGCGGAT	784
QY	781	CTATTTGGGAGCGGACGACGAGAGAGCGGAGCGCCGCAACGGGGCCCTGTGGGGACA	840
Db	785	CTATTTGGGAGCGGACGACGAGAGAGCGGAGCGCCGCAACGGGGCCCTGTGGGGACA	844
QY	841	GTGATCATCGGGAGCAAGAGCACCGTATAGACCTGCACATGATCCGGCTTACATGAA	900
Db	845	GTGATCATCGGGAGCAAGAGCACCGTATAGACCTGCACATGATCCGGCTTACATGAA	904
QY	901	GTGGTCACCCACGAGGGGTACTACGGCGAAGGCTTCAACGCCCATCATCTCTTCGAGCC	960
Db	905	GTGGTCACCCACGAGGGGTACTACGGCGAAGGCTTCAACGCCCATCATCTCTTCGAGCC	964
QY	961	TGCTTCTCTTCAGACAGAGCTCTCCCGACTACCTACATCATATGAGAACTCTTCTCG	1020
Db	965	TGCTTCTCTTCAGACAGAGCTCTCCCGACTACCTACATCATATGAGAACTCTTCTCG	1024
QY	1021	TAGCTCATCAGCAGCTTAGAGCTCTCTGGTGGAGCTTACATGATCGGTACCTGAC	1080
Db	1025	TAGCTCATCAGCAGCTTAGAGCTCTCTGGTGGAGCTTACATGATCGGTACCTGAC	1084

QY	1081	GGTGCACGCCCGCGGAGGATGCTTGGAAATCGGCTGGCTGAAGAGTGCTTACCAGATG	1140
Db	1085	GGTGCACGCCCGCGGAGGATGCTTGGAAATCGGCTGGCTGAAGAGTGCTTACCAGATG	1144
QY	1141	ATCGACCCGAGGTTGCGGAAAAAACTGAAAGTCTTTGATCATCGTCCAGCTCGTGGTTC	1200
Db	1145	ATCGACCCGAGGTTGCGGAAAAAACTGAAAGTCTTTGATCATCGTCCAGCTCGTGGTTC	1204
QY	1201	ATTGGGACTGTGTGGCCCATCTCTCGCCCTTTTCAATCAGGCTCAAGTTTCAACAAGATC	1260
Db	1205	ATTGGGACTGTGTGGCCCATCTCTCGCCCTTTTCAATCAGGCTCAAGTTTCAACAAGATC	1264
QY	1261	CAGTACGTGCACAGCTTGGAAAGACCTGGAGCAACTCATCTTATGGAACAAGTCCAGATC	1320
Db	1265	CAGTACGTGCACAGCTTGGAAAGACCTGGAGCAACTCATCTTATGGAACAAGTCCAGATC	1324
QY	1321	CCAGACTCGCTCTGCAATAACGAGAGAAAGACTGAAAGGCCAGGAGGAGCGCGAGG	1380
Db	1325	CCAGACTCGCTCTGCAATAACGAGAGAAAGACTGAAAGGCCAGGAGGAGCGCGAGG	1384
QY	1381	CCCAGCCGAGTGTGTCTGCCAGGCTCTGAAGAGAGCCAGAGGTGGCACCTAGTGAA	1440
Db	1385	CCCAGCCGAGTGTGTCTGCCAGGCTCTGAAGAGAGCCAGAGGTGGCACCTAGTGAA	1444
QY	1441	AACAGTCTGTCTGCTCTCAGAAAGATCAGGAAACCAAGCATGTCTGAGGGCAGCTGAGC	1500
Db	1445	AACAGTCTGTCTGCTCTCAGAAAGATCAGGAAACCAAGCATGTCTGAGGGCAGCTGAGC	1504
QY	1501	ATAACAAAGGACATGGAAGAAAGATTCAGATGCGCAAGAAACCTCTGTGAGAGCCCACTG	1560
Db	1505	ATAACAAAGGACATGGAAGAAAGATTCAGATGCGCAAGAAACCTCTGTGAGAGCCCACTG	1564
QY	1561	GCCGAGATCTCATCTCGCTCATCTGAGTCCCAATCTTCCAGGGTGGCCAGCCCTCC	1620
Db	1565	GCCGAGATCTCATCTCGCTCATCTGAGTCCCAATCTTCCAGGGTGGCCAGCCCTCC	1624
QY	1621	GTTCATCTCTGAAACCCAGCATCTTTTTCAGCTGTCTGAAACATTTGATTTTTTTTTT	1680
Db	1625	GTTCATCTCTGAAACCCAGCATCTTTTTCAGCTGTCTGAAACATTTGATTTT	1682
QY	1681	TAAAGATGAGTATTTGTGCGTTCCAGAAAGGGCCAGCTCTGAGCCCTCACCTTCC	1740
Db	1683	TAAAGATGAGTATTTGTGCGTTCCAGAAAGGGCCAGCTCTGAGCCCTCACCTTCC	1742
QY	1741	ACACTCAGAACTCTCAGCCGAGGAGGACAGAGCGAGGGGTGGCCCGGTGGCGTC	1800
Db	1743	ACACTCAGAACTCTCAGCCGAGGAGGACAGAGCGAGGGGTGGCCCGGTGGCGTC	1802
QY	1801	GGTGGCTTCCGCTCTCGTCCGAGCCCTGTGTGCTCAGAGCTGGATACAGATTTCAAGACC	1860
Db	1803	GGTGGCTTCCGCTCTCGTCCGAGCCCTGTGTGCTCAGAGCTGGATACAGATTTCAAGACC	1862
QY	1861	CTTCTCTTGTGTCAACCCGCTCAGGTTGGAGCCACAGACCCACCGCCACCCCGGT	1920
Db	1863	CTTCTCTTGTGTCAACCCGCTCAGGTTGGAGCCACAGACCCACCGCCACCCCGGT	1922
QY	1921	GGGTCTGCGTCTTCTGTGCTTCCCTCCAGAAATGGGCTCAGACCTAGAGAGCTCA	1980
Db	1923	GGGTCTGCGTCTTCTGTGCTTCCCTCCAGAAATGGGCTCAGACCTAGAGAGCTCA	1982
QY	1981	ACCCCTCTATGAGGGCCACGCTCTGGGTAGCTCTCTGACCTCCGACCTTATGTCCAAAT	2040
Db	1983	ACCCCTCTATGAGGGCCACGCTCTGGGTAGCTCTCTGACCTCCGACCTTATGTCCAAAT	2042
QY	2041	TCAACCACTGTTTTCATTTGACCCCGCTTCTCGCTCATTAATGACACCCAGCTCC	2100
Db	2043	TCAACCACTGTTTTCATTTGACCCCGCTTCTCGCTCATTAATGACACCCAGCTCC	2102
QY	2101	TTTGAGAGGATCAGAGCCCATTTGCAAGAGAGCGCTGCCAACCTCTTGCTCTCG	2160
Db	2103	TTTGAGAGGATCAGAGCCCATTTGCAAGAGAGCGCTGCCAACCTCTTGCTCTCG	2162

```
QY 2161 ATTGCAAAATGACACCCAGTATCTAGAACATTTCTAAGCCCTTTAACTCAGATGTC 2220
Db 2163 ATTGCAAAATGACACCCAGTATCTAGAACATTTCTAAGCCCTTTAACTCAGATGTC 2222
QY 2221 AGCCACCGGGCAAAACCCCGTCAATACCTCCACCAAGGAATAGATATGTGGACCTCACT 2280
Db 2223 AGCCACCGGGCAAAACCCCGTCAATACCTCCACCAAGGAATAGATATGTGGACCTCACT 2282
QY 2281 GCTCCCCCAACCCAGCGTCAAGCTGGGACACGCGCAACGCTGTTCCGGGTTGGAACAGCAG 2340
Db 2283 GCTCCCCCAACCCAGCGTCAAGCTGGGACACGCGCAACGCTGTTCCGGGTTGGAACAGCAG 2342
QY 2341 AGCTCAGAACTGCTCTGAATAGGACAGCTAGCAGAGAGAGAGATACAGGATTCGG 2400
Db 2343 AGCTCAGAACTGCTCTGAATAGGACAGCTAGCAGAGAGAGAGATACAGGATTCGG 2402
QY 2401 GCGTTGAGTGTTCAGAACTCATTCCGGAAGATAAATCAAGTGCCTGGCCGACCCAC 2460
Db 2403 GCGTTGAGTGTTCAGAACTCATTCCGGAAGATAAATCAAGTGCCTGGCCGACCCAC 2462
QY 2461 CTGCAATCAAGCTTGGACAGCGGGTCTTGTTCGGGAGGCAAAATTTCCCTAGGAAAAA 2520
Db 2463 CTGCAATCAAGCTTGGACAGCGGGTCTTGTTCGGGAGGCAAAATTTCCCTAGGAAAAA 2522
QY 2521 GAAGACAGACTTTTCTAATGTGTCCAAATGCGGATCACTGGTCAAGTGGACTCTAGAG 2580
Db 2523 GAAGACAGACTTTTCTAATGTGTCCAAATGCGGATCACTGGTCAAGTGGACTCTAGAG 2582
QY 2581 CACTGAGCTCCCTGTCTCTGGAAGTATTAAGAAAGAGCTGGCCAGGCAAGTGGCTCA 2640
Db 2583 CACTGAGCTCCCTGTCTCTGGAAGTATTAAGAAAGAGCTGGCCAGGCAAGTGGCTCA 2642
QY 2641 CGCTCTGTAATCCAGACTTTGGGAGGCGGAGGCGAGCGGATCACTGAGTGGAGGTTT 2700
Db 2643 CGCTCTGTAATCCAGACTTTGGGAGGCGGAGGCGGAGCGGATCACTGAGTGGAGGTTT 2702
QY 2701 GAGAACAGCTGGCCCAACATGTGAAACCTCTCTCTACTAATAAATACAAAAATAGCCA 2760
Db 2703 GAGAACAGCTGGCCCAACATGTGAAACCTCTCTCTACTAATAAATACAAAAATAGCCA 2762
QY 2761 GCGTGTGGCAGGTCCTGTAATCCAGTACTTGGGAGGCTGGGAGGATGAGATCACT 2820
Db 2763 GCGTGTGGCAGGTCCTGTAATCCAGTACTTGGGAGGCTGGGAGGATGAGATCACT 2822
QY 2821 TAAACCTGAGAGCAGAGGTTCAGTGGAGCAGAGTGGCCAGTCTGCACTTCCAGCTGG 2880
Db 2823 TAAACCTGAGAGCAGAGGTTCAGTGGAGCAGAGTGGCCAGTCTGCACTTCCAGCTGG 2882
QY 2881 CGACAGAGCAGACTCTGTCTCAAAAAAATAAAAAA 2917
Db 2883 CGACAGAGCAGACTCTGTCTCAAAAAAATAAAAAA 2919

RESULT 3
LOCUS BD183464
DEFINITION Novel genes and proteins encoded by the genes.
ACCESSION BD183464
VERSION BD183464.1 GI:31875664
KEYWORDS JP 2002345492-A/177.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 4976)
REFERENCE
AUTHORS Ohara, O., Nagase, T. and Nakajima, D.
TITLE Novel genes and proteins encoded by the genes.
JOURNAL Patent: JP 2002345492-A 177 03-DEC-2002;
KAZUSA DNA RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002345492-A/177
PD 03-DEC-2002
```

```
PF 26-FEB-2002 JP 2002049009
PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
PC C12N15/09, C07K14/47//A61K31/711, A61K38/00, A61K48/00, A61P25/00,
PC A61P25/14,
PC A61P25/18, A61P35/00, C12N15/00, A61K37/02
CC Novel genes and proteins encoded by the genes FH Key
FT CDS Location/Qualifiers
(404)..(1489).
FEATURES
source
1..4976
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.3%; Score 2896.6; DB 6; Length 4976;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2911; Conservative 0; Mismatches 4; Indels 2; Gaps 1;
QY 1 GCCGAGCCTCTGCCAGCCTGAGCTGGGAAGAGCAGCTACCTCGGAGGACGGCGCGCA 60
Db 5 GCCGAGCCTCTGCCAGCCTGAGCTGGGAAGAGCAGCTACCTCGGAGGACGGCGCGCA 64
QY 61 GCGCGGCGGATGAGAGGGGCGCAGCGCAGCCCCCGCTGGGGAGGCCACCGCTAAC 120
Db 65 GCGCGGCGGATGAGAGGGGCGCAGCGCAGCCCCCGCTGGGGAGGCCACCGCTAAC 124
QY 121 CTTGCACCCACCCACCCCTGCAAAAAGAGCTGGCGGGCGCTGGCCAGTGGCCCTGG 180
Db 125 CTTGCACCCACCCACCCCTGCAAAAAGAGCTGGCGGGCGCTGGCCAGTGGCCCTGG 184
QY 181 TGACCTTCTCGGATGAGAGTCCGCCCTCGGAGCATCTCTTCTCTCTAGGCTCTGAA 240
Db 185 TGACCTTCTCGGATGAGAGTCCGCCCTCGGAGCATCTCTTCTCTCTAGGCTCTGAA 244
QY 241 GCGCGGAGGCGGTGAGCGATGCCAGCTGCAACCGGCGCAGGCTTCGCTTGTGTCGA 300
Db 245 GCGCGGAGGCGGTGAGCGATGCCAGCTGCAACCGGCGCAGGCTTCGCTTGTGTCGA 304
QY 301 GTAAGGAGAGAGGCTGTCTGAGCTGAGAGGGGTTCCTCTTCAAGCAGTGCCTC 360
Db 305 GTAAGGAGAGAGGCTGTCTGAGCTGAGAGGGGTTCATCCCTGCTTCAAGCAGTGCCTC 364
QY 361 TTCCAGCTCCCATGGGACACCGAAGCAGCTCCCGATGCGAAGAGAGAGCGGGGTGAACTG 420
Db 365 TTCCAGCTCCCATGGGACACCGAAGCAGCTCCCGATGCGAAGAGAGAGCGGGGTGAACTG 424
QY 421 GAGGAATGCGAGGACGAAAGATCTTCCAGGGCCACTCCAGAGAGAGAGCGGGGTGAACTG 480
Db 425 GAGGAATGCGAGGACGAAAGATCTTCCAGGGCCACTCCAGAGAGAGAGCGGGGTGAACTG 484
QY 481 CTTGGCAGCCCGGTGGAAAGACACATCTCTCTCTCCAAACGCTAAATTTCAACGAGCG 540
Db 485 CTTGGCAGCCCGGTGGAAAGACACATCTCTCTCTCCAAACGCTAAATTTCAACGAGCG 544
QY 541 CATCTAAGAGAGAGAGCGCTGGTGGCCCGCAGAGATCAACATTTCTCTGATCAGAGTGA 600
Db 545 CATCTAAGAGAGAGAGCGCTGGTGGCCCGCAGAGATCAACATTTCTCTGATCAGAGTGA 604
QY 601 GGGTCCCTGCTGTCCGATGACTTTCTTGGATATCCCTGATGACCTGGATATTAACCTGGAT 660
Db 605 GGGTCCCTGCTGTCCGATGACTTTCTTGGATATCCCTGATGACCTGGATATTAACCTGGAT 664
QY 661 GACATCGAGAGACCCCGATGAGACCGAATCTCGCTGGAGTTCTCTGGGAAATGGCAACGACTG 720
Db 665 GACATCGAGAGACCCCGATGAGACCGAATCTCGCTGGAGTTCTCTGGGAAATGGCAACGACTG 724
QY 721 GAGTGGGAGAGAGACACCCCGTGGCCACCCCGCAAGAAATATCCCGGGGAGACGCGCGGAT 780
Db 725 GAGTGGGAGAGAGACACCCCGTGGCCACCCCGCAAGAAATATCCCGGGGAGACGCGCGGAT 784
QY 781 CTATTGTTGGGAGCGGCACGACGAGGAGCGCGCCCAACGGCGCTCTGTGGCGGACA 840
```



Db 785 CTATTTGGGACGGCAGCAGGAGGAGGCGCGCCGCCAAACGGGGCCCTGTGTGGCGGACA 844  
Qy 841 GTGATCATCGGGAGCAAGAGCAGCCGTATAGACCTGCAACATGATCCGGCTTACATGAATA 900  
Db 845 GTGATCATCGGGAGCAAGAGCAGCCGTATAGACCTGCAACATGATCCGGCTTACATGAATA 904  
Qy 901 GTGGTCACCCACGGAGGGTACTAGCGGGAAGGCTCAACGCCATCATCTGCTTCGAGCC 960  
Db 905 GTGGTCACCCACGGAGGGTACTAGCGGGAAGGCTCAACGCCATCATCTGCTTCGAGCC 964  
Qy 961 TGCTTCCTCCACAGCAGCCTCCCGCACTACCACTACATCATGAGAACCTCTTCCTG 1020  
Db 965 TGCTTCCTCCACAGCAGCCTCCCGCACTACCACTACATCATGAGAACCTCTTCCTG 1024  
Qy 1021 TACGTATCATCAGCAGCTTAGAGCTCTCGTGGCTGAGGACTACATGATCGTGTACCTGAAC 1080  
Db 1025 TACGTATCATCAGCAGCTTAGAGCTCTCGTGGCTGAGGACTACATGATCGTGTACCTGAAC 1084  
Qy 1081 GGTGCCACGCCCGGGAGGATGCTGGAATCGGCTGGTGAAGAGTGTCTACAGATG 1140  
Db 1085 GGTGCCACGCCCGGGAGGATGCTGGAATCGGCTGGTGAAGAGTGTCTACAGATG 1144  
Qy 1141 ATCGACCGGAGTTCGGGAAAACCTGAAGTCTTCATCATGTCACCCCTCGTGTTC 1200  
Db 1145 ATCGACCGGAGTTCGGGAAAACCTGAAGTCTTCATCATGTCACCCCTCGTGTTC 1204  
Qy 1201 ATTCCGACTGTGTGGCCATCTCTCGCCCTTTCATCAGGGTCAAGTTCATCAACAGATC 1260  
Db 1205 ATTCCGACTGTGTGGCCATCTCTCGCCCTTTCATCAGGGTCAAGTTCATCAACAGATC 1264  
Qy 1261 CAGTAGCTGCAAGCTTGGAGAACCTGGAGCAACTCATCCCTATGAAACAGTTCAGATC 1320  
Db 1265 CAGTAGCTGCAAGCTTGGAGAACCTGGAGCAACTCATCCCTATGAAACAGTTCAGATC 1324  
Qy 1321 CCAGATCGCTCTGCAATACGAAGAGGAGGAGTGAAGGCGGAGGAGGAGGCGGAG 1380  
Db 1325 CCAGATCGCTCTGCAATACGAAGAGGAGGAGTGAAGGCGGAGGAGGAGGCGGAG 1384  
Qy 1381 CCCAGCCGGAGTTTGTGCTGCCAGGTCTGAAGAGAAGCCAGAGGTGCAACAGTGGAA 1440  
Db 1385 CCCAGCCGGAGTTTGTGCTGCCAGGTCTGAAGAGAAGCCAGAGGTGCAACAGTGGAA 1444  
Qy 1441 AACAGTCTGCTGTGCTCAGAAATCAGGAAACAGCATGTCTGAGGCGAGGTGAGC 1500  
Db 1445 AACAGTCTGCTGTGCTCAGAAATCAGGAAACAGCATGTCTGAGGCGAGGTGAGC 1504  
Qy 1501 ATAAACAGGACATGAGAGATCCAGATGCGCAGAAACCTCTCTCAGAGCCCACTG 1560  
Db 1505 ATAAACAGGACATGAGAGATCCAGATGCGCAGAAACCTCTCTCAGAGCCCACTG 1564  
Qy 1561 GCCCAGATCTCATCTCGCTCATCTCTGAGTCCCAATCTTCAAGGGTGCAGCCCTCC 1620  
Db 1565 GCCCAGATCTCATCTCGCTCATCTCTGAGTCCCAATCTTCAAGGGTGCAGCCCTCC 1624  
Qy 1621 GTTCATCTGAAACCCAGATCTCTTTTTCAGTGTGTTGAAACATTTATTTTTTTTT 1680  
Db 1625 GTTCATCTGAAACCCAGATCTCTTTTTCAGTGTGTTGAAACATTTATTTTTTTTT 1682  
Qy 1681 TAACGATGAGTATTTGTGGTTCCAGAAAGGGCCAGCTCTGAGCCCTCAACCTTCC 1740  
Db 1683 TAACGATGAGTATTTGTGGTTCCAGAAAGGGCCAGCTCTGAGCCCTCAACCTTCC 1742  
Qy 1741 ACACTACGAACTCTCAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800  
Db 1743 ACACTACGAACTCTCAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1802  
Qy 1801 GGTGGCTCGCTCTGCTGAGCCCTCTGCTGAGAGTGTGATCAAGATTCAGAGC 1860  
Db 1803 GGTGGCTCGCTCTGCTGAGCCCTCTGCTGAGAGTGTGATCAAGATTCAGAGC 1862  
Qy 1861 CTCTCTCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920  
Db 1863 CTCTCTCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1922

RESULT 4  
AB058775  
LOCUS

Qy 1921 GGGTCTGGTCTCTTCTGTGCTTTCCCTCCAGAAATCGGCTCAGACCTAGAAAGCTCA 1980  
Db 1923 GGGTCTGGTCTCTTCTGTGCTTTCCCTCCAGAAATCGGCTCAGACCTAGAAAGCTCA 1982  
Qy 1981 ACCCCCTATGAGGGCCACGTCTCGGGTAGTCTCTGACCTCCGACCTTATGTCCAAAT 2040  
Db 1983 ACCCCCTATGAGGGCCACGTCTCGGGTAGTCTCTGACCTCCGACCTTATGTCCAAAT 2042  
Qy 2041 TCACACCCCATGTTTTCATTTGACCCGCCCTTCTCGTCATATGACACCCAGCTCC 2100  
Db 2043 TCACACCCCATGTTTTCATTTGACCCGCCCTTCTCGTCATATGACACCCAGCTCC 2102  
Qy 2101 TTTGAGAGGATCAGAGCCCATTTGCAAGAAGAGCGCTGCCAACCATCTTGTCTCCG 2160  
Db 2103 TTTGAGAGGATCAGAGCCCATTTGCAAGAAGAGCGCTGCCAACCATCTTGTCTCCG 2162  
Qy 2161 ATTGAAAATGACACCCCAAGTAATCTGAAACATTTCTCAAGGCCCTTTTAACTCAGATGCA 2220  
Db 2163 ATTGAAAATGACACCCCAAGTAATCTGAAACATTTCTCAAGGCCCTTTTAACTCAGATGCA 2222  
Qy 2221 AGCCACCGGGCAAAACCCCGTCAATCTCCCAACCAAGGAAATGAGATATGTGGACCTCACT 2280  
Db 2223 AGCCACCGGGCAAAACCCCGTCAATCTCCCAACCAAGGAAATGAGATATGTGGACCTCACT 2282  
Qy 2281 GCTCCCCCAACCCAGCGTCAGGCTGGGACACAGCCAAACGCTTCCGGGTTGGAAACAGCAG 2340  
Db 2283 GCTCCCCCAACCCAGCGTCAGGCTGGGACACAGCCAAACGCTTCCGGGTTGGAAACAGCAG 2342  
Qy 2341 AGGCTCAGAAACTGTGCTCTGAAATAGGACAGCTTAGCAAGAGGAAAGATACAGGGTATCGG 2400  
Db 2343 AGGCTCAGAAACTGTGCTCGGAAATAGGACAGCTTAGCAAGAGGAAAGATACAGGGTATCGG 2402  
Qy 2401 GGGTTGAGTCTTTTCAGAAAGTCAATTCGGGAAAGATAAATCCAGTGGCTGGCGCAGCCAC 2460  
Db 2403 GGGTTGAGTCTTTTCAGAAAGTCAATTCGGGAAAGATAAATCCAGTGGCTGGCGCAGCCAC 2462  
Qy 2461 CTGCATTCAGAACTTGGACACGCGGTTCTTTGTTCCGGAGGCAAAATTTCCCTAGGAAAAA 2520  
Db 2463 CTGCATTCAGAACTTGGACACGCGGTTCTTTGTTCCGGAGGCAAAATTTCCCTAGGAAAAA 2522  
Qy 2521 GAAGACAGACTTTTCTTAATGTGGTCCAAATGCGGATCACTGGTCAAGTGGACTCTAGAG 2580  
Db 2523 GAAGACAGACTTTTCTTAATGTGGTCCAAATGCGGATCACTGGTCAAGTGGACTCTAGAG 2582  
Qy 2581 CACTGAGCTCCCTGTCTCTGGAATTTTAAAGAAAGGCTGGCCAGGACGATGGCTCA 2640  
Db 2583 CACTGAGCTCCCTGTCTCTGGAATTTTAAAGAAAGGCTGGCCAGGACGATGGCTCA 2642  
Qy 2641 CGCCTGTAAATCCAGACTTTTGGAGCCGAGCGAGCGGATCACTGAGGTGAGGAGTTT 2700  
Db 2643 CGCCTGTAAATCCAGACTTTTGGAGCCGAGCGAGCGGATCACTGAGGTGAGGAGTTT 2702  
Qy 2701 GAGAAACAGCTGGCCAAACATGTGAAACCTCACTCTACTAAAAATACAAAATTAGCCA 2760  
Db 2703 GAGAAACAGCTGGCCAAACATGTGAAACCTCACTCTACTAAAAATACAAAATTAGCCA 2762  
Qy 2761 GGGCTGTGGCAGGTGCTTAATCCAGCTACTTGGGAGGCTGAGGATGAGATCACT 2820  
Db 2763 GGGCTGTGGCAGGTGCTTAATCCAGCTACTTGGGAGGCTGAGGATGAGATCACT 2822  
Qy 2821 TAAACCTCAGAGGCGAGGTTACAGTGAAGCAAGATCGTGGCAGCTGCAATCCAGCGCTGG 2880  
Db 2823 TAAACCTCAGAGGCGAGGTTACAGTGAAGCAAGATCGTGGCAGCTGCAATCCAGCGCTGG 2882  
Qy 2881 CGACAGAGAGAGCTCTGTCTCAAAAAAATAAAAAA 2917  
Db 2883 CGACAGAGAGAGCTCTGTCTCAAAAAAATAAAAAA 2919

AB058775 4976 bp mRNA linear PRI 05-JUN-2001

```

DEFINITION Homo sapiens mRNA for KIAA1872 protein, partial cds.
ACCESSION AB058775
VERSION 1
KEYWORDS GI:14017960
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
Hominidae; Homo.
REFERENCE Nagase, T., Nakayama, M., Nakajima, D., Kikuno, R. and Ohara, O.
AUTHORS Prediction of the coding sequences of unidentified human genes. XX.
TITLE The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro
JOURNAL DNA Res. 8 (2), 85-95 (2001)
PUBMED 11347906
REFERENCE 2 (bases 1 to 4976)
AUTHORS Ohara, O., Nagase, T. and Kikuno, R.
TITLE Direct Submission
JOURNAL Submitted (27-MAR-2001) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail: cdnainfo@kazusa.or.jp,
URL: http://www.kazusa.or.jp/huge, Tel: 81-438-52-3913,
Fax: 81-438-52-3914)
FEATURES
    Location/Qualifiers
        1..4976
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="bj05518"
            /tissue_type="brain"
            /note="vector: pBluescriptII SK plus"
        1..4976
            /gene="KIAA1872"
            <188..1492
            /gene="KIAA1872"
            /note="Start codon is not identified."
            /codon_start=1
            /product="KIAA1872 protein"
            /protein_id="BAB47501.1"
            /db_xref="GI:14017961"
            /translation="PSSDARSAPASILFLGSEGPSVSDAQLHFRARLCLPVRRRG
            CLSCGVIIPASQCLFPAPMGTEATLRNENVDVKEWODELPLRPLFTGTGVELLGS
            PVEDTSSPTLNFAGHRRKRLTVAPEINISLQSEGLSDFLDPTDIDINVDI
            IETPDTSLEFLGNLEWEDDTPVATKNMGDSADLFGDGTTEGDSAAAGRLWR
            TWIIGQEHRIIDLHMRPYMKVTHGGVYEGEGLNLIIVFAACFLPDSLSLPDHYTIMEN
            LFLVYSSLELLAVEDYIMVILNGATPRRMPGICWKLKCYQIMIDRLRLKNLXSLIV
            HPSMRTVLAISRPPISVKPIFKTYVHSLDLSQLIPMERHVQIPDCVLTVEERLK
            ARRESARPOPEFLPRSEKPEVAPENRSALVSDQETSMS"
ORIGIN
Query Match 99.3%; Score 2896.6; DB 8; Length 4976;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2911; Conservative 0; Mismatches 4; Indels 2; Gaps 1;
QY 1 GCGAGCCTCTGCGAGCTGAGCTGGGAAGAGCAGTACCTCGGAGGCGGCGCA 60
DB 5 GCGAGCCTCTGCGAGCCTGAGCTGGGAAGAGCAGTACCTCGGAGGCGGCGCA 64
QY 61 GCGGCGCGGATGAGAGGGGCGCGCGAGCGCGCGCTGGGAGCCACCGCTAAC 120
DB 65 GCGGCGCGGATGAGAGGGGCGCGAGCGCGCGCTGGGAGCCACCGCTAAC 124
QY 121 CTGTACACCCACCCCTGACAAAGAGCTGGGCGGCTGGGCGGCTGGGCGGCTGG 180
DB 125 CTGTACACCCACCCCTGACAAAGAGCTGGGCGGCTGGGCGGCTGGGCGGCTGG 184
QY 181 TGACCTTCTCGGATGAGAGTCCGCGCTGCGAGCATCTCTCTCTCTAGGCTCTGAA 240
DB 185 TGACCTTCTCGGATGAGAGTCCGCGCTGCGAGCATCTCTCTCTCTAGGCTCTGAA 244
QY 241 GCGCGGCGGAGCGTGGAGCGATGCCAGCTGCACCGCGGAGGCGCTCGCTTTGTTGCCA 300

```

```

DB 245 GCGCGGCGGAGCGTGGAGCGATGCCAGCTGCACCGCGGCGAGGCTCGCTTTGTTGCCA 304
QY 301 GTAAGGAGGAGAGGCTGTCTCAGCTGCAGAGGGGTCACTCCCTGCTTCAAGCCAGTGCCTC 360
DB 305 GTAAGGAGGAGAGGCTGTCTCAGCTGCAGAGGGGTCACTCCCTGCTTCAAGCCAGTGCCTC 364
QY 361 TTCCAGCTCCCATCGGGGACCAACCGAAGCCACGCTCCGGATGGAAAAACGTGACGCTGAAG 420
DB 365 TTCCAGCTCCCATCGGGGACCAACCGAAGCCACGCTCCGGATGGAAAAACGTGACGCTGAAG 424
QY 421 GAGGAATGGCAGGAGGAGATCTTCCAGGCCCATCTCCAGAAAGAGACGCGGGGTGGAACGTG 480
DB 425 GAGGAATGGCAGGAGGAGATCTTCCAGGCCCATCTCCAGAAAGAGACGCGGGGTGGAACGTG 484
QY 481 CTGTGCAGCCCGGTGGAAAGACACATCTCTCTCCCTCCAAACACGCTAAATTTCAACCGAGCG 540
DB 485 CTGTGCAGCCCGGTGGAAAGACACATCTCTCTCCCTCCAAACACGCTAAATTTCAACCGAGCG 544
QY 541 CATCTAAGAGGAAGACGCTGTGTGGCCCCCAGAGATCAACATTTCTCTGGATCAGAGTGAAG 600
DB 545 CATCTAAGAGGAAGACGCTGTGTGGCCCCCAGAGATCAACATTTCTCTGGATCAGAGTGAAG 604
QY 601 GGGTCCCTGCTGTCCGATGACCTTCTTGATATACCCCTGATGACCTGGATTAATTAACGTGAT 660
DB 605 GGGTCCCTGCTGTCCGATGACCTTCTTGATATACCCCTGATGACCTGGATTAATTAACGTGAT 664
QY 661 GACATCGAGACCCCGATGAGACCGACTCGCTGGAGTTCTTGGGGAATGGCAACGACTG 720
DB 665 GACATCGAGACCCCGATGAGACCGACTCGCTGGAGTTCTTGGGGAATGGCAACGACTG 724
QY 721 GAGTGGGAAGACGACACCCCGTGGCCACCGCCAAAGAAACATGCCCGGGGACGACGCGGAT 780
DB 725 GAGTGGGAAGACGACACCCCGTGGCCACCGCCAAAGAAACATGCCCGGGGACGACGCGGAT 784
QY 781 CTATTTGGGAGCGGACGACGAGCGAGGACCGGCGCGCCGCAACGCGGCGCTGTGCGGAC 840
DB 785 CTATTTGGGAGCGGACGACGAGCGAGGACCGGCGCGCCGCAACGCGGCGCTGTGCGGAC 844
QY 841 GTGATCATCGGGGAGCAGACCGTATAGACCTGACATGATCGGAGCTTACATGATAA 900
DB 845 GTGATCATCGGGGAGCAGACCGTATAGACCTGACATGATCGGAGCTTACATGATAA 904
QY 901 GTGCTCACCCACGAGGGGTACTAGCGGCAAGCGCTCAACGGCCATCATGCTTCCAGCC 960
DB 905 GTGCTCACCCACGAGGGGTACTAGCGGCAAGCGCTCAACGGCCATCATGCTTCCAGCC 964
QY 961 TGCTTCTTCCAGACGAGCGCTCCCGCACTACCACTACATCATGAGAAACCTTTCCTG 1020
DB 965 TGCTTCTTCCAGACGAGCGCTCCCGCACTACCACTACATCATGAGAAACCTTTCCTG 1024
QY 1021 TACGTTCATCAGCAGCTTAGAGCTCTGTGTGGCTGAGGACTACATGATCGTGTACCTGAAC 1080
DB 1025 TACGTTCATCAGCAGCTTAGAGCTCTGTGTGGCTGAGGACTACATGATCGTGTACCTGAAC 1084
QY 1081 GGTGCCAGCCCGCGGAGGATGCTGGAATCGGCTGCTGGAAGAGTGTCTACCAAGATG 1140
DB 1085 GGTGCCAGCCCGCGGAGGATGCTGGAATCGGCTGCTGGAAGAGTGTCTACCAAGATG 1144
QY 1141 ATGCAACGAGGTTGCGGAAAAAACCTGAAAGTCTTGTGATCATCGTCCACCCCTCGTGGTTC 1200
DB 1145 ATGCAACGAGGTTGCGGAAAAAACCTGAAAGTCTTGTGATCATCGTCCACCCCTCGTGGTTC 1204
QY 1201 ATTGGAGTGTGCTGGCCATCTCTGCGCTTTTCAATCAGGCTCAGTTTCAATCAAGATC 1260
DB 1205 ATTGGAGTGTGCTGGCCATCTCTGCGCTTTTCAATCAGGCTCAGTTTCAATCAAGATC 1264
QY 1261 CAGTACGTGACAGCTTGGAAAGACCTGGAGCAACTCATCCCTTATGGAAACGCTCCAGATC 1320
DB 1265 CAGTACGTGACAGCTTGGAAAGACCTGGAGCAACTCATCCCTTATGGAAACGCTCCAGATC 1324
QY 1321 CCAGACTGCGTCTTGCAATACGAAGAGGAAAGACTGAAGGCCAGGAGGAGGAGGCGGAGG 1380
DB 1325 CCAGACTGCGTCTTGCAATACGAAGAGGAAAGACTGAAGGCCAGGAGGAGGAGGCGGAGG 1384

```









Qy	968	TTCCAGACGACGAGCTCTCCCGACTACCACTACATCATGGAGAACTCTTCTCTGTACGTCA	1027
Db	960	TTCCAGACGACGAGCTCTCCCGACTACCACTACATCATGGAGAACTCTTCTCTGTACGTCA	1019
Qy	1028	TCAGCAGCTTAGAGCTCTGTGTGGCTGAGGACTACATGATCGTGTACTTGAAACGGTGCCA	1087
Db	1020	TCAGCAGCTTAGAGCTCTGTGTGGCTGAGGACTACATGATCGTGTACTTGAAACGGTGCCA	1079
Qy	1088	CGCCCCGGCGGAGGATCGCTGGAATCGGCTGGCTGGAAGAGTGCTACACAGATGATCGAAC	1147
Db	1080	CGCCCCGGCGGAGGATCGCTGGAATCGGCTGGCTGGAAGAGTGCTACACAGATGATCGAAC	1139
Qy	1148	GGAGGTTGCGGAAAAACCTGAAGTCTCTTGATCATCGTCCACCCCTCGTGTTCATTGCGA	1207
Db	1140	GGAGGTTGCGGAAAAACCTGAAGTCTCTTGATCATCGTCCACCCCTCGTGTTCATTGCGA	1199
Qy	1208	CTGTGCTGGCCATCTCTCGCCCTTTTCATCAGGTCRAAGTTTCATCAAACAAGATCGATPAG	1267
Db	1200	CTGTGCTGGCCATCTCTCGCCCTTTTCATCAGGTCRAAGTTTCATCAAACAAGATCGATPAG	1259
Qy	1268	TGCACAGCTTGGAAAGACTCGGAGCAACTCATCCCTATGGAACACGTCACAGATCCAGACT	1327
Db	1260	TGCACAGCTTGGAAAGACTCGGAGCAACTCATCCCTATGGAACACGTCACAGATCCAGACT	1319
Qy	1328	CGTCTCTGCAATA CGAAGAGAAAGACTGAAGCCAGAGGGAGAGCGCGAGGCCCCAGC	1387
Db	1320	CGTCTCTGCAATA CGAAGAGAAAGACTGAAGGCCAGAGGGAGAGCGCGAGGCCCCAGC	1379
Qy	1388	CGGAGTTGTGCTGCGCCAGGTCCTGAAGAGAACCGAGAGGTGGACACAGTGGAAAAACAGGT	1447
Db	1380	CGGAGTTGTGCTGCGCCAGGTCCTGAAGAGAACCGAGAGGTGGACACAGTGGAAAAACAGGC	1439
Qy	1448	CTGCTCTGGTCTCAGAAAGATCAGGAAACAAAGCATGCTCTGAGGCGAGCTGAGCATTAACAA	1507
Db	1440	CTGCTCTGGTCTCAGAAAGATCAGGAAACAAAGCATGCTCTGAGGCGAGCTGAGCATTAACAA	1499
Qy	1508	AGGACATGGAAGAAAGATTCCAGATGCGCAGAAAACTCTGTGACAGCCCACTGGCCCCAG	1567
Db	1500	AGGACATGGAAGAAAGATTCCAGATGCGCAGAAAACTCTGTGACAGGCCCACTGGCCCCAG	1559
Qy	1568	ATCTCATCTGCTCATCTGAGTCCCAATCTTCCAAAGGTTGCAGCCCTCGTTCATC	1627
Db	1560	ATCTCATCTGCTCATCTGAGTCCCAATCTTCCAAAGGTTGCAGCCCTCGTTCATC	1619
Qy	1628	TCTGAAAACCCAGCATCTTTTCAGTCTGCTTGAAAAACATTGATTTTTTTTTTTTAAACGAT	1687
Db	1620	TCTGAAAACCCAGCATCTTTTCAGTCTGCTTGAAAAACATTGATTTTTTTTTTTTAAACGAT	1679
Qy	1688	GCAGTATTGTGGTTTCAGAAAAAGGCCCGCAGCTCTGAGCCCTCAACCTTCACACTCA	1747
Db	1680	GCAGTATTGTGGTTTCAGAAAAAGGCCCGCAGCTCTGAGCCCTTCACACTCA	1739
Qy	1748	CGAACTCTCAGCCGAGGAAGGCAAGAGCCGAGGGGTGGCCCGTGGCGTGGTGCGC	1807
Db	1740	CGAACTCTCAGCCGAGGAAGGCAAGAGCCGAGGGGTGGCCCGTGGCGTGGTGCGC	1799
Qy	1808	TCGCTCTCTGCTCGCAGCCCTCTGTGTGTCAGAGCTTGGATACAAGATTTCAGACCCCTTCTCT	1867
Db	1800	TCGCTCTCTGCTCGCAGCCCTCTGTGTGTCAGAGCTTGGATACAAGATTTCAGACCCCTTCTCT	1859
Qy	1868	TGCTTGTACCCGCTCCAGGTTGGAGCCACAGACACCCACCGCCACCCCGGCTGGGCTCTG	1927
Db	1860	TGCTTGTACCCGCTCCAGGTTGGAGCCACAGACACCCACCGCCACCCCGGCTGGGCTCTG	1919
Qy	1928	CGTCTTTCTGTGCTTTTCCTCCAGATTGGGGCTTCAGACTAGAGCTCAACCCCGC	1987
Db	1920	CGTCTTTCTGTGCTTTTCCTCCAGATTGGGGCTTCAGACTAGAGCTCAACCCCGC	1979
Qy	1988	TATGAGGGCCACGTCCTGGGGTAGCTCTCACTCCGACCTTAATGTCCTCAAC	2047
Db	1980	TATGAGGGCCACGTCCTGGGGTAGCTCTCACTCCGACCTTAATGTCCTCAAC	2039
Qy	2048	CATGGTTTTTCATTTTGACCGCCGCCCTTCTCGCTCATATAATGACCCAGCTCTTTGAGA	2107

Db	2040		CATGGTTTTTCAATTTGACCGCGCCCTTCTCGCTCAATAATGACACCCAGCTCTTTGAGA	2099
Qy	2108	GGATCAGAGCCCATTTGCACAAGAGAGAGCCGCTGCCAACCATCTTCTCTCCGATTCGAA	2167	
Db	2100	GGATCAGAGCCCATTTGCACAAGAGAGAGCCGCTGCCAACCATCTTCTCTCCGATTCGAA	2159	
Qy	2168	AATGACACCCCGAGTAATCTAGAAACATTTCTCAAGCCCTTTAACTCAGATGTCAAGCCACC	2227	
Db	2160	AATGACACCCCGAGTAATCTAGAAACATTTCTCAAGCCCTTTAACTCAGATGTCAAGCCACC	2219	
Qy	2228	GGGCAAAACCCGCTCAATACCTCCCAACAAGGAATGAGATATGFGGACCTCACTGTCTCCC	2287	
Db	2220	GGGCAAAACCCGCTCAATACCTCCCAACAAGGAATGAGATATGFGGACCTCACTGTCTCCC	2279	
Qy	2288	CAACCCAGCGTCAGGCTGGGACACGCGCAACGCTGTTCCGGTTGGAAACAGCAGAGGCTCA	2347	
Db	2280	CAACCCAGCGTCAGGCTGGGACACGCGCAACGCTGTTCCGGTTGGAAACAGCAGAGGCTCA	2339	
Qy	2348	GAATACTGGCTCTGAAATAGGACAGACCTAGCAAGAGGAAGATACAGGGTATCGGGCGTTG	2407	
Db	2340	GAATACTGGCTCTGAAATAGGACAGACCTAGCAAGAGGAAGATACAGGGTATCGGGCGTTG	2399	
Qy	2408	AGTGTTCAGAGTCAATTCGGGAAGATAAATCCAGTGCCTGGCCGACGCCACCTGCAATT	2467	
Db	2400	AGTGTTCAGAGTCAATTCGGGAAGATAAATCCAGTGCCTGGCCGACGCCACCTGCAATT	2459	
Qy	2468	CAGAAGCTTGACACGCGGTTCTGTTCCGGAGGCAAAATTCCTAGGAAAGAAAGACA	2527	
Db	2460	CAGAAGCTTGACACGCGGTTCTGTTCCGGAGGCAAAATTCCTAGGAAAGAAAGACA	2519	
Qy	2528	GACTTTTCTAATGTGGTCCAAATGCGGATCACATGGTTCAGATGAGCTCTAGAGACGCTGAG	2587	
Db	2520	GACTTTTCTAATGTGGTCCAAATGCGGATCACATGGTTCAGATGAGCTCTAGAGACGCTGAG	2579	
Qy	2588	CTCCCTGTCTCTGGAAGTATTTAAGAAAAAGCTGGGCCAGGCAACGATGGCTCACGCCCTGT	2647	
Db	2580	CTCCCTGTCTCTGGAAGTATTTAAGAAAAAGCTGGGCCAGGCAACGATGGCTCACGCCCTGT	2639	
Qy	2648	AATCCAGACTTTGGGAGCGCGAGGCGGATCACTTGAGGTGAGGATTTGAGAAACA	2707	
Db	2640	AATCCAGACTTTGGGAGGTGCGAGGCGGATCACTTGAGGTGAGGATTTGAGAAACA	2699	
Qy	2708	GCCTGGCCACATGGTGAACCTCATCTCTACTTAAATAATACAAAATATAGCCAGCGGTGG	2767	
Db	2700	GCCTGGCCACATGGTGAACCTCATCTCTACTTAAATAATACAAAATATAGCCAGCGGTGG	2759	
Qy	2768	TGGCAGGTGCTGTAAATCCAGCTACTTGGGAGGCTGAGGCATGAGAAATCACTTAAACCT	2827	
Db	2760	TGGCAGGTGCTGTAAATCCAGCTACTTGGGAGGCTGAGGCATGAGAAATCACTTAAACCT	2819	
Qy	2828	GAGAGGAGAGGTTACAGTGAGCCCAAGATCTGTGCCACTGCAATCCAGGCTGGGCGACAGA	2887	
Db	2820	GAGAGGAGAGGTTACAGTGAGCCCAAGATCTGTGCCACTGCAATCCAGGCTGGGCGACAGA	2879	
Qy	2888	GCAAGACTCTGTCTC	2902	
Db	2880	GCAAGACTCTGTCTC	2894	
RESULT 7				
LOCUS	AX881045		2894 bp	DNA
DEFINITION	Sequence 15950 from Patent EP1074617.		linear	PAT 17-DEC-2003
ACCESSION	AX881045			
VERSION	AX881045.1			
KEYWORDS	GI:40035781			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
	Hominidae; Homo.			
REFERENCE	1			

LOCUS AX881045 2894 bp DNA  
DEFINITION Sequence 15950 from Patent Epi074617.  
linear PAT 17-DEC-2003

ACCESSION AX881045  
VERSION AX881045.1 GI:40035781

**KEYWORDS**  
**SOURCE** . Homo sapiens (human)

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

## REFERENCE

AUTHORS Ota,T., Isoqai,T., Nishikawa,T., Hayaishi,K., Saito,K., Yamamoto,J.,  
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
TITLE Primers for synthesising full-length cDNA and their use  
JOURNAL Patent: EP 1074617-A 15950 07-FEB-2001;  
Research Association for Biotechnology (JP)

## FEATURES

## source

1. 2894  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
872..1480  
/note="unnamed protein product"

## CDS

/codon\_start=1  
/protein\_id="CAB91182.1"  
/db\_xref="GI:40035782"  
/translation="MIRPYMKVTHGGYEGEGLNATIVPAACFLPDSSLPDYHYIMEN  
LFLVTSLELLVAEDYMIYVLNGATPRRMPGICHLKCYQVIDRLRLKNUKSLIIV  
HESWFRTVLASRPISVKSINKIQYVHSLDLEQLIMEHVQIPDCVLQVEERLKL  
ARRSARPOPEFLPRSEBKPEVAPENRPAIVSDQETSMS"

## ORIGIN

Query Match 98.4%; Score 2870.2; DB 6; Length 2894;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2886; Conservative 0; Mismatches 8; Indels 1; Gaps 1;  
QY 8 CTCTGCAGCCCTGAGCTGGAGAAAGACGACTACCTCGGAGGCGAGGCGCGAGGCGGC 67  
DB 1 CTCTGCAGCCCGGGCTGGGAAGAGCAGCTACCTCGGAGGCGAGGCGCGAGGCGGC 60  
QY 68 GCGATGAGAGGGGGCGAGCCCGAGCCCGCGCTGGGAGGCCACCGCTAACCCCTGCAC 127  
DB 61 GCGATGAGAGGGGGCGAGCCCGAGCCCGCGCTGGGAGGCCACCGCTAACCCCTGCAC 120  
QY 128 CCCACCCACCCCTGTCACAAAGAGCTGGGGGGCTGGCCACGCTCGCCCTGGGTGACCTT 187  
DB 121 CCCACCCACCCCTGTCACAAAGAGCTGGGGGGCTGGCCACGCTCGCCCTGGGTGACCTT 180  
QY 188 CCTCGATGAGAAATCGCCCTTCGAGAGCATCTCTTCCTCTAGGCTCTGAAGGCCCGG 247  
DB 181 CCTCGATGAGAAATCGCCCTTCGAGAGCATCTCTTCCTCTAGGCTCTGAAGGCCCGG 240  
QY 248 GAGCGTGAAGGATGCCAGCTGCAACCGGAGGCGCTGCTTGTGTCGAGTAAGGA 307  
DB 241 GAGCGTGAAGGATGCCAGCTGCAACCGGAGGCGCTGCTTGTGTCGAGTAAGGA 300  
QY 308 GAGAGGCTGTCTCAGCTGAGAGGGGTCACTCCCTCTCAAGCCAGTGCCTCTCCAG 367  
DB 301 GAGAGGCTGTCTCAGCTGAGAGGGGTCACTCCCTCTCAAGCCAGTGCCTCTCCAG 360  
QY 368 CTCCCATGGGACCAACGAGCCACGCTCCGAGTGGAAACGTTGAGCGTGAAGGAGGAAT 427  
DB 361 CTCCCATGGGACCAACGAGCCACGCTCCGAGTGGAAACGTTGAGCGTGAAGGAGGAAT 420  
QY 428 GCGAGACGAGATCTTCCAGGCGCACTCCAGAGAGAGCGGGGTGGAACTGCTTGGCA 487  
DB 421 GCGAGACGAGATCTTCCAGGCGCACTCCAGAGAGAGCGGGGTGGAACTGCTTGGCA 480  
QY 488 GCCCGGTGGAAGACACATCTCTCTCCCAACGCTTAAATTTCAACGGAGCGCATCGTA 547  
DB 481 GCCCGGTGGAAGACACATCTCTCTCCCAACGCTTAAATTTCAACGGAGCGCATCGTA 540  
QY 548 AGAGGAAGCGCTGGTGGGCCCCAGAGATCAACATTTCTCTGGATCAGAGTGAAGGCTCC 607  
DB 541 AGAGGAAGCGCTGGTGGGCCCCAGAGATCAACATTTCTCTGGATCAGAGTGAAGGCTCC 600  
QY 608 TGCTGTCCGATGACTTCTTGGATACCCCTGATGACCTGGATATTAACGTGATGATCG 667  
DB 601 TGCTGTCCGATGACTTCTTGGATACCCCTGATGACCTGGATATTAACGTGATGATCG 660  
QY 668 AGACCCCGGATGAGACCGACTCGCTCGAGTTCCTGGGGAATGGCAACGAACCTGGAGTGG 727  
DB 661 AGACCCCGGATGAGACCGACTCGCTCGAGTTCCTGGGGAATGGCAACGAACCTGGAGTGG 720

QY 728 AAGACGACACCCCGCTGGCCACCGCCCAAGAAACATGCCCCGGGACAGCGCGATCTATTTC 787  
DB 721 AAGACGACACCCCGCTGGCCACCGCCCAAGAAACATGCCCCGGGACAGCGCGATCTATTTC 780  
QY 788 GGGACGGACACGACGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 847  
DB 781 GGGACGGACACGACGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840  
QY 848 TCGGGGAGCAGAGGACCGGTATAGACCTGCAATGATCCGGCTTACATGAAGTGGTCA 907  
DB 841 TCGGGGAGC-AGAGGACCGGTATAGACCTGCAATGATCCGGCTTACATGAAGTGGTCA 899  
QY 908 CCCACGAGGGTACTACGCGGAGGCGCTCAAGGCCATCATGCTTTCGAGCGCTGCTTCC 967  
DB 900 CCCACGAGGGTACTACGCGGAGGCGCTCAAGGCCATCATGCTTTCGAGCGCTGCTTCC 959  
QY 968 TTTCCAGACAGCGACCTCCCGGACACATCATCATCATCATCATCATCATCATCATCATCAT 1027  
DB 960 TTTCCAGACAGCGACCTCCCGGACACATCATCATCATCATCATCATCATCATCATCATCAT 1019  
QY 1028 TCAGCAGCTTACAGCTCTCTGGTGGCTGAGGACTACATGATCGTGACCTGAAACGGTGCCA 1087  
DB 1020 TCAGCAGCTTACAGCTCTCTGGTGGCTGAGGACTACATGATCGTGACCTGAAACGGTGCCA 1079  
QY 1088 CGCCCCGGCGGAGGATGCTCGGAATCGGCTGGCTGAAGAAGTGTCTACAGATGATCGGACC 1147  
DB 1080 CGCCCCGGCGGAGGATGCTCGGAATCGGCTGGCTGAAGAAGTGTCTACAGATGATCGGACC 1139  
QY 1148 GGAGGTTGCGGGAACCACTGAAAGTCTTGTATCATCGTCCACCCCTCGTGGTTCATTCGGA 1207  
DB 1140 GGAGGTTGCGGGAACCACTGAAAGTCTTGTATCATCGTCCACCCCTCGTGGTTCATTCGGA 1199  
QY 1208 CTGTCTCGCCCATCTCTCGCCCTTTCATCAGCGTCAAGTTCATCAACAGATCCAGTAGC 1267  
DB 1200 CTGTCTCGCCCATCTCTCGCCCTTTCATCAGCGTCAAGTTCATCAACAGATCCAGTAGC 1259  
QY 1268 TGCACAGCTTGGGAAGACCTGGAGCAACTCATCTCTATGGAAACACCTCCAGATCCAGACT 1327  
DB 1260 TGCACAGCTTGGGAAGACCTGGAGCAACTCATCTCTATGGAAACACCTCCAGATCCAGACT 1319  
QY 1328 GGGTCTCTGAATACGAAGAGGAAAGTGAAGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGG 1387  
DB 1320 GGGTCTCTGAATACGAAGAGGAAAGTGAAGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGG 1379  
QY 1388 GGGAGTGTGTGCTGCCAGGCTCTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1447  
DB 1380 GGGAGTGTGTGCTGCCAGGCTCTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1439  
QY 1448 CTGCTCTGCTCAGGAAGATCAGGAACCAAGCATGTCTCTGAGGGGACGCTGAGCATAACAA 1507  
DB 1440 CTGCTCTGCTCAGGAAGATCAGGAACCAAGCATGTCTCTGAGGGGACGCTGAGCATAACAA 1499  
QY 1508 AGGACATGGAAGAGATTCAGAGATGCCAGAAAACTCTGTCTGACGCGCCACTGCGGCCGAG 1567  
DB 1500 AGGACATGGAAGAGATTCAGAGATGCCAGAAAACTCTGTCTGACGCGCCACTGCGGCCGAG 1559  
QY 1568 ATCTCATCTGCTCATCTGAGTCCCAATCTTCCAGGGGTGCGAGCCCTCCGTTTCATC 1627  
DB 1560 ATCTCATCTGCTCATCTGAGTCCCAATCTTCCAGGGGTGCGAGCCCTCCGTTTCATC 1619  
QY 1628 TCTGAAACCCAGCATCTCTTTCAGCTGCTTGAACCAATTTGATTTTTTTTTTTTAAACGAT 1687  
DB 1620 TCTGAAACCCAGCATCTCTTTCAGCTGCTTGAACCAATTTGATTTTTTTTTTTTAAACGAT 1679  
QY 1688 GCAGATTTTGTGCGTTCCAGAAAAAGGGCCAGCTCTGAGCCCTCAACCTTCCACATCA 1747  
DB 1680 GCAGATTTTGTGCGTTCCAGAAAAAGGGCCAGCTCTGAGCCCTCAACCTTCCACATCA 1739  
QY 1748 CGAACTCTCAGCCGAGGAGGCAAGAGCGAGGGGGTGGCCGCGCTGCGGTGGCC 1807  
DB 1740 CGAACTCTCAGCCGAGGAGGCAAGAGCGAGGGGGTGGCCGCGCTGCGGTGGCC 1799  
QY 1808 TCCGCTCTGCTCGCAGGCCCTGTGGTTCAGAGCTGGATCAAGATTCAAGACCTTCTCT 1867

Db 1800 TCGCTCTGCTCGACCCCTGTGGTCAGAGCTGGATCAAGATTCAGACCCCTTCTCT 1859  
Qy 1868 TGCTGTGTCACCGCTCCAGGTGTGAGCCACAGACACCCACCGCCCGCTGGGTCTG 1927  
Db 1860 TGCTGTGTCACCGCTCCAGGTGTGAGCCACAGACACCCACCGCCCGCTGGGTCTG 1919  
Qy 1928 CGTCTCTTCTGTCGCTTCTGTCAGAAATGCGGCTTCAGACCTAGAGCTCAACCCCCC 1987  
Db 1920 CGTCTCTTCTGTCGCTTCTGTCAGAAATGCGGCTTCAGACCTAGAGCTCAACCCCCC 1979  
Qy 1988 TATGAGGCCACGCTCCGCGGTAGCTCTGACCTCCGACCTTATGTCCTCAATTTTCACAC 2047  
Db 1980 TATGAGGCCACGCTCCGCGGTAGCTCTGACCTCCGACCTTATGTCCTCAATTTTCACAC 2039  
Qy 2048 CATGTTTCTTTCATTTGACCCCGCCCTTCTCGCTCATATGACACCCAGCTCTTTGAGA 2107  
Db 2040 CATGTTTCTTTCATTTGACCCCGCCCTTCTCGCTCATATGACACCCAGCTCTTTGAGA 2099  
Qy 2108 GATCAGAGCCCATTTGCAAGAGAGCGCTGCAACCATCTTGTCTCCGATTCGAA 2167  
Db 2100 GATCAGAGCCCATTTGCAAGAGAGCGCTGCAACCATCTTGTCTCCGATTCGAA 2159  
Qy 2168 AATGACACCCCGCTTCTAGACATTTCTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2227  
Db 2160 AATGACACCCCGCTTCTAGACATTTCTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2219  
Qy 2228 GGCACAAACCCCGCTCAATACCTCCACCAAGGAATGATATGGAACCTCCTGCTCCCG 2287  
Db 2220 GGCACAAACCCCGCTCAATACCTCCACCAAGGAATGATATGGAACCTCCTGCTCCCG 2279  
Qy 2288 CAACCCAGCGTCAAGCTGGGACACGCAACGCTGTTCCGGTGGGAACAGCAGGCTCA 2347  
Db 2280 CAACCCAGCGTCAAGCTGGGACACGCAACGCTGTTCCGGTGGGAACAGCAGGCTCA 2339  
Qy 2348 GAACCTGGCTCTGAATAGGACACCTAGCAAGAGGAATACAGGGTATCGGGCTTTG 2407  
Db 2340 GAACCTGGCTCTGAATAGGACACCTAGCAAGAGGAATACAGGGTATCGGGCTTTG 2399  
Qy 2408 AGTGTCTTCAAGCTCAATCGGGAAGATAAATCAGTCCGCTGCGCCAGCCACTCTGAT 2467  
Db 2400 AGTGTCTTCAAGCTCAATCGGGAAGATAAATCAGTCCGCTGCGCCAGCCACTCTGAT 2459  
Qy 2468 CAAAGCTTGGACACGAGGGTCTTGTTCGGGAGGCAATTTCCCTAGGAAAAGAGACA 2527  
Db 2460 CAAAGCTTGGACACGAGGGTCTTGTTCGGGAGGCAATTTCCCTAGGAAAAGAGACA 2519  
Qy 2528 GACTTTCTTAATGTTGTCMAATGCGGATCACTGGTCAGATGACTCTAGAGCACTGAG 2587  
Db 2520 GACTTTCTTAATGTTGTCMAATGCGGATCACTGGTCAGATGACTCTAGAGCACTGAG 2579  
Qy 2588 CTCCTCTCTCTGGAAGTATTAAGAAAAGGCTGGCCAGGACGATGCTCAAGCTGT 2647  
Db 2580 CTCCTCTCTCTGGAAGTATTAAGAAAAGGCTGGCCAGGACGATGCTCAAGCTGT 2639  
Qy 2648 AATCCAGACTTTGGGAGGCCAGGACGCGGATCACTGAGTGAAGGAGTTTGAACA 2707  
Db 2640 AATCCAGACTTTGGGAGGCCAGGACGCGGATCACTGAGTGAAGGAGTTTGAACA 2699  
Qy 2708 GCTTGCCCAACATGTTGAAACCTCATCTTCTAATAAATAACAAAATTAGCCAGCGGTG 2767  
Db 2700 GCTTGCCCAACATGTTGAAACCTCATCTTCTAATAAATAACAAAATTAGCCAGCGGTG 2759  
Qy 2768 TGSCAGTCTCTGTAATCCAGCTACTTGGAGGCTGAGGCATGAGATCACTTAACCT 2827  
Db 2760 TGSCAGTCTCTGTAATCCAGCTACTTGGAGGCTGAGGCATGAGATCACTTAACCT 2819  
Qy 2828 GAGAGCAGAGGTTTACAGTGAGCCAGATCTGTCACCTCATTTCCAGCTGCGGCACAGA 2887  
Db 2820 GAGAGCAGAGGTTTACAGTGAGCCAGATCTGTCACCTCATTTCCAGCTGCGGCACAGA 2879  
Qy 2888 GCAAGACTCTGTCT 2902

Db 2880 GCAAGACTCTGTCT 2894

## RESULT 8

AK027889

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

2894 bp mRNA linear PRI 30-JAN-2004  
Homo sapiens cDNA FLJ14983 fis, clone Y79AA1000268, weakly similar  
to Mus musculus Nip21 mRNA.  
AK027889  
AK027889.1 GI:14042891  
oligo capping; fis (full insert sequence).  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

REFERENCE

AUTHORS

TITLE

JOURNAL





Db 1920 CGTCCTTCTGTCCTTCCCTCCAGATGCGGCTCAGACCTAGAGCTCAACCCCC 1979  
QY 1988 TATGAGGCGCAGCTCTGCGGTAGCTCTGACCTCCGACCTTAGTCCAAATTTACACC 2047  
Db 1980 TATGAGGCGCAGCTCTGCGGTAGCTCTGACCTCCGACCTTAGTCCAAATTTACACC 2039  
QY 2048 CATGGTTTTCATTTGACCGCCCTTCTGCTCTATATGACACCGAGCTCTTTGAGA 2107  
Db 2040 CATGGTTTTCATTTGACCGCCCTTCTGCTCTATATGACACCGAGCTCTTTGAGA 2099  
QY 2108 GATCAGAGCCCATTTGACAGAGAGCGCTCCCAACCATCTTCTGCTCCGATTCGAA 2167  
Db 2100 GATCAGAGCCCATTTGACAGAGAGCGCTCCCAACCATCTTCTGCTCCGATTCGAA 2159  
QY 2168 AATGACACCCAGTAATCTAGAACATTTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2227  
Db 2160 AATGACACCCAGTAATCTAGAACATTTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2219  
QY 2228 GGGCAACCCCGTCAATACCTCCCAAGGAATGAGATATGTGGACCTCACTGTCCCC 2287  
Db 2220 GGGCAACCCCGTCAATACCTCCCAAGGAATGAGATATGTGGACCTCACTGTCCCC 2279  
QY 2288 CAACCCAGGCTCAGGCTGGGACACGCAACGCTGTTCCGGTTGGACAGCAGGCTCA 2347  
Db 2280 CAACCCAGGCTCAGGCTGGGACACGCAACGCTGTTCCGGTTGGACAGCAGGCTCA 2339  
QY 2348 GAACTGGCTCTGAAATAGGCGAGACCTAGCAAGAGGAAGATACAGGCTATCGGGTTTG 2407  
Db 2340 GAACTGGCTCTGAAATAGGCGAGACCTAGCAAGAGGAAGATACAGGCTATCGGGTTTG 2399  
QY 2408 AGTGTTCAGAGTCAATTCGGGAAGATAATCCAGTGCCTGGCGCACCCACCTGCAAT 2467  
Db 2400 AGTGTTCAGAGTCAATTCGGGAAGATAATCCAGTGCCTGGCGCACCCACCTGCAAT 2459  
QY 2468 CAAAGCTTCGACAGCGGCTTCTGTTCCGGAGGCAATTTCCCTAGGAAAAGAGACA 2527  
Db 2460 CAAAGCTTCGACAGCGGCTTCTGTTCCGGAGGCAATTTCCCTAGGAAAAGAGACA 2519  
QY 2528 GACTTTCTTAATGTGTCCAAATGCGGATCACTGGTTCAGATGAGTCTTAGAGCACTGAG 2587  
Db 2520 GACTTTCTTAATGTGTCCAAATGCGGATCACTGGTTCAGATGAGTCTTAGAGCACTGAG 2579  
QY 2588 CTCCTGTCTCTGGAAGTATTTAAGAAAGGCTGGGCCAGGACGATGCTCACGCTGT 2647  
Db 2580 CTCCTGTCTCTGGAAGTATTTAAGAAAGGCTGGGCCAGGACGATGCTCACGCTGT 2639  
QY 2648 AATCCAGACTTTGGGAGGCCAGGACGCGGATCAGCTGAGGTGAGGATTTGAGAAC 2707  
Db 2640 AATCCAGACTTTGGGAGGTCGAGGCGGCGGATCACCTGAGGTGAGGATTTGAGAAC 2699  
QY 2708 GCCTGGCCAAATGGTGAACCTCTATCTTAAATAACAAATAGCCAGGCTGG 2767  
Db 2700 GCCTGGCCAAATGGTGAACCTCTATCTTAAATAACAAATAGCCAGGCTGG 2759  
QY 2768 TGGCAGGTGCTGTATATCCAGCTACTTGGGAGGCTGAGGATGAGATCACTTAAACT 2827  
Db 2760 TGGCAGGTGCTGTATATCCAGCTACTTGGGAGGCTGAGGATGAGATCACTTAAACT 2819  
QY 2828 GAGAGCGAGGTTTACAGTGAGCCAGATGCTGCCACTGCAATTCAGCTGGGCGACAGA 2887  
Db 2820 GAGAGCGAGGTTTACAGTGAGCCAGATGCTGCCACTGCAATTCAGCTGGGCGACAGA 2879  
QY 2888 GCAAGACTCTGCTC 2902  
Db 2880 GCAAGACTCTGCTC 2894

## RESULT 9

AX747472

LOCUS

DEFINITION

AX747472

ACCESSION

AX747472.1

VERSION

Sequence 997 from Patent EP1308459.

AX747472

AX747472.1

GI:32131860

2759 bp

mRNA

linear

PAT 20-JUN-2003

## KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

AUTHORS

Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuno, Y.

## TITLE

JOURNAL

FEATURES

source

1..2759

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

ORIGIN

Query Match

Best Local Similarity

Matches 2756; Conservative

0; Mismatches

3; Indels

136; Gaps

2;

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY	728	AAGACGACACCCCGTGGCCACCGCCCAAGAA	CATGCCCGGGGACAGCGCGGATCTATT	787
DB	585	AAGACGACACCCCGTGGCCACCGCCCAAGAA	CATGCCCGGGGACAGCGCGGATCTATT	644
QY	788	GGGACGGCAGCAGGAGGACGGCAGCGCCCAAG	CGGCGCCCTGTGGCGGACAGGATCA	847
DB	645	GGGACGGCAGCAGGAGGACGGCAGCGCCCAAG	CGGCGCCCTGTGGCGGACAGGATCA	704
QY	848	TGGGGAGCAAGAGCACCGTATAGACCTGCACAT	GCATCGGCCCTTACATGAAGTGGTCA	907
DB	705	TGGGGAGCAAGAGCACCGTATAGACCTGCACAT	GCATCGGCCCTTACATGAAGTGGTCA	764
QY	908	CCGACGGAGGTACTAGCGGAGAGGCTCAAGCC	CAATCATCTGTCTTCGAGCGCTGCTCC	967
DB	765	CCGACGGAGGTACTAGCGGAGAGGCTCAAGCC	CAATCATCTGTCTTCGAGCGCTGCTCC	824
QY	968	TTCCAGACAGCAGCTCCCGGACTACCACTACAT	CATCATGAGAACCTCTTCGTAGCTCA	1027
DB	825	TTCCAGACAGCAGCTCCCGGACTACCACTACAT	CATCATGAGAACCTCTTCGTAGCTCA	884
QY	1028	TCAGCAGCTTAGAGCTCTGTGGCTGAGGACTAC	ATGATCGTGTACCTGAAACGGTGC	1087
DB	885	TCAGCAGCTTAGAGCTCTGTGGCTGAGGACTAC	ATGATCGTGTACCTGAAACGGTGC	944
QY	1088	CGCCCGCGGAGGATGCTGTGAATCGGCTGCGT	GAAGTGTCTACCAAGATGATCGAC	1147
DB	945	CGCCCGCGGAGGATGCTGTGAATCGGCTGCGT	GAAGTGTCTACCAAGATGATCGAC	1004
QY	1148	GGAGGTTGCGGAAAACCTGAACTGCTTGTAT	CATCGTCCACCGCTGTGTTTCAATCGGA	1207
DB	1005	GGAGGTTGCGGAAAACCTGAACTGCTTGTAT	CATCGTCCACCGCTGTGTTTCAATCGGA	1064
QY	1208	CTGTGCTGCGCATCTCTGCGCTTTCATCAGG	GTCAAGTTCATCAACAGATCCAGTACG	1267
DB	1065	CTGTGCTGCGCATCTCTGCGCTTTCATCAGG	GTCAAGTTCATCAACAGATCCAGTACG	1124
QY	1268	TGCACAGCTTGGAGACCTGGAGCAACTCAT	CCCTATGGAACAGTCCAGACT	1327
DB	1125	TGCACAGCTTGGAGACCTGGAGCAACTCAT	CCCTATGGAACAGTCCAGACT	1184
QY	1328	GGTCTGTCATACGAGAGGAAAGACTGAAGC	CAAGGAGGAGCGCGGAGCCCGACG	1387
DB	1185	GGTCTGTCATACGAGAGGAAAGACTGAAGC	CAAGGAGGAGCGCGGAGCCCGACG	1244
QY	1388	CGGAGTTTGTGTGCTGCCAGTCTGAAGAG	AGCCAGAGGTGCACAGTGGAAACAGGT	1447
DB	1245	CGGAGTTTGTGTGCTGCCAGTCTGAAGAG	AGCCAGAGGTGCACAGTGGAAACAGGT	1304
QY	1448	CTGCTCTGCTCTCAGAAAGATCAGGAAACA	AGCATGTCTGAGGCGACGTGAGCATAACA	1507
DB	1305	CTGCTCTGCTCTCAGAAAGATCAGGAAACA	AGCATGTCTGAGGCGACGTGAGCATAACA	1364
QY	1508	AGGACATGGAAGAGATTCCAGATGCGAGAA	ACCTCTGTCTAGAGCGCCACCTGGCCCGAG	1567
DB	1365	AGGACATGGAAGAGATTCCAGATGCGAGAA	ACCTCTGTCTAGAGCGCCACCTGGCCCGAG	1424
QY	1568	ATCTCATCTGCTCATCTGAGTCCCAATCTT	CCAAGGTGCCAGCCCTCGTTTCATC	1627
DB	1425	ATCTCATCTGCTCATCTGAGTCCCAATCTT	CCAAGGTGCCAGCCCTCGTTTCATC	1484
QY	1628	TCTGAAACCCAGCATCTCTTTCAGCTGCTT	GAAACATTTGTTTTTTTTTTTAAACGAT	1687
DB	1485	TCTGAAACCCAGCATCTCTTTCAGCTGCTT	GAAACATTTGTTTTTTTTTTTAAACGAT	1544
QY	1688	GCAGTATTTGTGGTTTCCAGAAAGGGCCAG	CTCTGAGCCCTCACTTCCACACTCA	1747
DB	1545	GCAGTATTTGTGGTTTCCAGAAAGGGCCAG	CTCTGAGCCCTCACTTCCACACTCA	1604
QY	1748	CGAATCTCAGCGGAGGAGGAGGAGGAGGAG	GAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1807
DB	1605	CGAATCTCAGCGGAGGAGGAGGAGGAGGAG	GAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1664

QY	1808	TCCGCTCTCTGCTCGAGCCCTCTGTGGTCA	GAGCTGGATACAAAGATTCAAGACCTTCTCT	1867
DB	1665	TCCGCTCTCTGCTCGAGCCCTCTGTGGTCA	GAGCTGGATACAAAGATTCAAGACCTTCTCT	1724
QY	1868	TGCTTGTCAACCGCTTCCAGGTTTGGAGCC	ACAGACCCACCGCGCTGGGTCTG	1927
DB	1725	TGCTTGTCAACCGCTTCCAGGTTTGGAGCC	ACAGACCCACCGCGCTGGGTCTG	1784
QY	1928	CGTCTTCTCTGTGCTTTTCCCTCCAGAAT	CGGCTCAGACCTAGAAAGCTCAACCCGCC	1987
DB	1785	CGTCTTCTCTGTGCTTTTCCCTCCAGAAT	CGGCTCAGACCTAGAAAGCTCAACCCGCC	1844
QY	1988	TATGAGGCGCAGCTCTGGGGTAGCTCTGAC	CTCCGACCTTATGTCCAAATTTTCAAC	2047
DB	1845	TATGAGGCGCAGCTCTGGGGTAGCTCTGAC	CTCCGACCTTATGTCCAAATTTTCAAC	1904
QY	2048	CATGCTTTTTCATTTGACCCCGCTTCTCG	CTCATTAATGACCCAGCTCTTTTGAGA	2107
DB	1905	CATGCTTTTTCATTTGACCCCGCTTCTCG	CTCATTAATGACCCAGCTCTTTTGAGA	1964
QY	2108	GGATCAGAGCCCATTTGCAAAAGAGCGCT	GGCCCAACCTCTTGTCTCCGATTGCAA	2167
DB	1965	GGATCAGAGCCCATTTGCAAAAGAGCGCT	GGCCCAACCTCTTGTCTCCGATTGCAA	2024
QY	2168	AATGACACCCAGTAATCTAGAAATTTCT	CAAGCCCTTTTAACTCAGATGTCAAGC	2227
DB	2025	AATGACACCCAGTAATCTAGAAATTTCT	CAAGCCCTTTTAACTCAGATGTCAAGC	2084
QY	2228	GGGCAACCCCGTCAATACCTCCCAAGAG	AGGATATGTGGACCTCACTGCTCC	2287
DB	2085	GGGCAACCCCGTCAATACCTCCCAAGAG	AGGATATGTGGACCTCACTGCTCC	2144
QY	2288	CAACCCAGCTCAGCTGGGACACGCCAA	CCCTGTTCCGGTTGGAAAGAGGCTCA	2347
DB	2145	CAACCCAGCTCAGCTGGGACATGCCAAC	CCCTGTTCCGGTTGGAAAGAGGCTCA	2204
QY	2348	GAACCTGCTCTGAAATAGGAGACCTAG	CAAGAGGAGATACAGGGTATCGGGCTTTG	2407
DB	2205	GAACCTGCTCTGAAATAGGAGACCTAG	CAAGAGGAGATACAGGGTATCGGGCTTTG	2264
QY	2408	AGTGTTCAGAACTCCTCGGAGATATAAT	CCAGTGGCTGGCGGACCACTGCAAT	2467
DB	2265	AGTGTTCAGAACTCCTCGGAGATATAAT	CCAGTGGCTGGCGGACCACTGCAAT	2324
QY	2468	CAAGCTTGGACACGCGGTTCTTGTTCG	GAGGCAAAATTTCCCTAGGAAAGAGACA	2527
DB	2325	CAAGCTTGGACACGCGGTTCTTGTTCG	GAGGCAAAATTTCCCTAGGAAAGAGACA	2384
QY	2528	GACTTTTCTAATGTGGTCCAAATGGGAT	CACTGGTCHAGATGGACTTAGAAGCACTGAG	2587
DB	2385	GACTTTTCTAATGTGGTCCAAATGGGAT	CACTGGTCHAGATGGACTTAGAAGCACTGAG	2444
QY	2588	CTCCCTGTCTGTGAGATTTTAAAGAA	AGGCTGGGCCAGGACATGGCTCACGCTGT	2647
DB	2445	CTCCCTGTCTGTGAGATTTTAAAGAA	AGGCTGGGCCAGGACATGGCTCACGCTGT	2504
QY	2648	AATCCAGACTTTGGAGCGCGGAGCGAG	CGGATCACCTGAGGTGAGGATTTTGAACA	2707
DB	2505	AATCCAGACTTTGGAGCGCGGAGCGAG	CGGATCACCTGAGGTGAGGATTTTGAACA	2564
QY	2708	GCCTGGCCAAACATGTGTGAACCTCAT	CTCTACTATAAAATACAAAATTTAGCCAG	2767
DB	2565	GCCTGGCCAAACATGTGTGAACCTCAT	CTCTACTATAAAATACAAAATTTAGCCAG	2624
QY	2768	TGGCAGGTGCTGTGTAATCCAGCTTA	CTTGGAGGCTGAGGATGAGATCACTTAA	2827
DB	2625	TGGCAGGTGCTGTGTAATCCAGCTTA	CTTGGAGGCTGAGGATGAGATCACTTAA	2684
QY	2828	GAGAGGAGAGGTTACGTGAGCCAGAT	CTGTCACCTGCACTTCAGCTGGCCAGCA	2887
DB	2685	GAGAGGAGAGGTTACGTGAGCCAGAT	CTGTCACCTGCACTTCAGCTGGCCAGCA	2744
QY	2888	GCAAGACTCTCTGCTC	2902	





Db 465 TGCTGTCGATGACTTCTTTGGATACCCCTGATGACCTGGATATTAAAGTGGATGATCG 524  
Qy 668 AGACCCCGATGAGACCGACTCGCTGGAGTTCTGGGAAATGGCAACGAACCTGGAGTGGG 727  
Db 525 AGACCCCGATGAGACCGACTCGCTGGAGTTCTGGGAAATGGCAACGAACCTGGAGTGGG 584  
Qy 728 AAGACGACACCCCGTGGCCACCGCAAGAACATGCGCGGGGACAGCGGCGATCTATTGG 787  
Db 585 AAGACGACACCCCGTGGCCACCGCAAGAACATGCGCGGGGACAGCGGCGATCTATTGG 644  
Qy 788 GGGACGCGACGACGAGGAGCGGACGCGCGCCCAACGCGCGCTGTGCGCGACAGTATCA 847  
Db 645 GGGACGCGACGACGAGGAGCGGACGCGCGCCCAACGCGCGCTGTGCGCGACAGTATCA 704  
Qy 848 TCGGGGAGCAAGAGCACCGTATAGACTGACATGATCGGCGCTTACATGAAAGTGGTCA 907  
Db 705 TCGGGGAGCAAGAGCACCGTATAGACTGACATGATCGGCGCTTACATGAAAGTGGTCA 764  
Qy 908 CCCACGGAGGTTACTAGCGGAAAGGCTCAA GCGCATCATGCTTTTCGAGCGTGTCC 967  
Db 765 CCCACGGAGGTTACTAGCGGAAAGGCTCAA GCGCATCATGCTTTTCGAGCGTGTCC 824  
Qy 968 TTCACGACGACGCTCCCGACTACCACTACATCATGAGAACCTTCTCTGTAGCTCA 1027  
Db 825 TTCACGACGACGCTCCCGACTACCACTACATCATGAGAACCTTCTCTGTAGCTCA 884  
Qy 1028 TCAGCAGCTTAGAGCTCTGCTGGCTGAGGACTACATGATGCTGTAACCTGAAACGCTGCCA 1087  
Db 885 TCAGCAGCTTAGAGCTCTGCTGGCTGAGGACTACATGATGCTGTAACCTGAAACGCTGCCA 944  
Qy 1088 CGCCCGGGGAGGATGCTGGAAATCGCTGGCTGGAAGAGTGTACACAGATGATCGACC 1147  
Db 945 CGCCCGGGGAGGATGCTGGAAATCGCTGGCTGGAAGAGTGTACACAGATGATCGACC 1004  
Qy 1148 GGAAGTTGCGGAAACCTGAACTGCTGATCGTCCACCCCTCGTGGTTCATTGCGA 1207  
Db 1005 GGAAGTTGCGGAAACCTGAACTGCTGATCATGCTCACCCCTCGTGGTTCATTGCGA 1064  
Qy 1208 TGTGCTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTTCATCAACAGATCCAGTAGC 1267  
Db 1065 TGTGCTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTTCATCAACAGATCCAGTAGC 1124  
Qy 1268 TGCAAGCTTGGAAACCTGGAGCAACTCATCCTTATGAAACAGTGTCCAGACT 1327  
Db 1125 TGCAAGCTTGGAAACCTGGAGCAACTCATCCTTATGAAACAGTGTCCAGACT 1184  
Qy 1328 GGTCTCTCAATACGAGGAGGAAAGTGAAGCCAGGAGGAGCGCGAGCCCGCCAGC 1387  
Db 1185 GGTCTCTCAATACGAGGAGGAAAGTGAAGCCAGGAGGAGCGCGAGCCCGCCAGC 1244  
Qy 1388 CGGAGTTTGTGCTGCTGAGTCTGAAGAGGAGGAGGAGTGGCAGGAGTGGCAACAGGAGT 1447  
Db 1245 CGGAGTTTGTGCTGCTGAGTCTGAAGAGGAGGAGGAGTGGCAGGAGTGGCAACAGGAGT 1304  
Qy 1448 CTGCTCTGCTCTCAGAAAGTACAGAAACAGCATGCTCTGAGCGCAGCTGAGCATAACAA 1507  
Db 1305 CTGCTCTGCTCTCAGAAAGTACAGAAACAGCATGCTCTGAGCGCAGCTGAGCATAACAA 1364  
Qy 1508 AGGACATGAGAGAGGATTCAGATGCGCAGAAACCTCTGTGAGCGCCACTGCGCCCGAG 1567  
Db 1365 AGGACATGAGAGAGGATTCAGATGCGCAGAAACCTCTGTGAGCGCCACTGCGCCCGAG 1424  
Qy 1568 ATCTCATCTGCTCATCTGAGTCCCAATCTTCCAAAGGAGTGGCAGGCGCTCGGTTCAATC 1627  
Db 1425 ATCTCATCTGCTCATCTGAGTCCCAATCTTCCAAAGGAGTGGCAGGCGCTCGGTTCAATC 1484  
Qy 1628 TCTGAAACCCAGCATCTCTTTTTCAGCTGTGAAACATTTGTAATTTTTTTTTTTTAAACGAT 1687  
Db 1485 TCTGAAACCCAGCATCTCTTTTTCAGCTGTGAAACATTTGTAATTTTTTTTTTTTAAACGAT 1544  
Qy 1688 GCAGTATTTGTGCTTCCAGAAAGGCGGCGCTCTGAGCCCTCACCCCTTCCACACTCA 1747  
Db 1545 GCAGTATTTGTGCTTCCAGAAAGGCGGCGCTCTGAGCCCTCACCCCTTCCACACTCA 1604

Qy 1748 CGAACTCTCAGCGGAGGAGCAAGAACGACGAGGGGTGGCCCGGTGGCTGGCTGGTGGCC 1807  
Db 1605 CGAACTCTCAGCGGAGGAGCAAGAACGACGAGGGGTGGCCCGGTGGCTGGTGGCC 1664  
Qy 1808 TCGCTCTCTGCTCGAGCGCTTGTGGTCAAGCTGGATACAGAAATTCAGAACCTTCTCT 1867  
Db 1665 TCGCTCTCTGCTCGAGCGCTTGTGGTCAAGCTGGATACAGAAATTCAGAACCTTCTCT 1724  
Qy 1868 TGTCTGTCAACCGCTCCAGGTTGGAGCCACAGACACCCACCGCCCGCTGGGTCTG 1927  
Db 1725 TGTCTGTCAACCGCTCCAGGTTGGAGCCACAGACACCCACCGCCCGCTGGGTCTG 1784  
Qy 1928 CGTCTCTTCTGCTGCTTTCCTCCAGAAATGGGCTCAGACCTAGAACCTCAACCCCGC 1987  
Db 1785 CGTCTTCTGCTGCTTTCCTCCAGAAATGGGCTCAGACCTAGAACCTCAACCCCGC 1844  
Qy 1988 TATGAGGCGCAGCTCTGCGGTGAGCTCTCGACCTCCGACCTTATGTCCAAAATTCACACC 2047  
Db 1845 TATGAGGCGCAGCTCTGCGGTGAGCTCTCGACCTCCGACCTTATGTCCAAAATTCACACC 1904  
Qy 2048 CATGCTTTTTCATTTGACCGCGCTTCTCGCTCATTAATGACCCAGCTCTCTTGAGA 2107  
Db 1905 CATGCTTTTTCATTTGACCGCGCTTCTCGCTCATTAATGACCCAGCTCTCTTGAGA 1964  
Qy 2108 GGATCAGAGCCCATTTGCAACAGAGCGCTGCGCAACCATCTTGTCTCCGATTGCAA 2167  
Db 1965 GGATCAGAGCCCATTTGCAACAGAGCGCTGCGCAACCATCTTGTCTCCGATTGCAA 2024  
Qy 2168 AATGACACCCAGTAAATCTAGAACATTTCTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2227  
Db 2025 AATGACACCCAGTAAATCTAGAACATTTCTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2084  
Qy 2228 GGGCAAAACCGCTGCAATACCTCCACCAAGAGATGAGATATGTGACCTCACTGCTCCCG 2287  
Db 2085 GGGCAAAACCGCTGCAATACCTCCACCAAGAGATGAGATATGTGACCTCACTGCTCCCG 2144  
Qy 2288 CAACCCAGCGTCAGGCTGGGACACCGCAACCGCTGTTCCGGGTGGAAACAGCAGAGGCTCA 2347  
Db 2145 CAACCCAGCGTCAGGCTGGGACATGCCAACGCTGTTCCGGGTGGAAACAGCAGAGGCTCA 2204  
Qy 2348 GAAACTGCTCTGAAATAGGCGAGCTAGCAAGAGAGAAATACAGGGTATCGGGCTTTG 2407  
Db 2205 GAAACTGCTCTGAAATAGGCGAGCTAGCAAGAGAGAAATACAGGGTATCGGGCTTTG 2264  
Qy 2408 AGTGTTCAGAGTCAATTCGGGAGATTAATTCAGTGGCTGGCCGCGCAGCTGCAAT 2467  
Db 2265 AGTGTTCAGAGTCAATTCGGGAGATTAATTCAGTGGCTGGCCGCGCAGCTGCAAT 2324  
Qy 2468 CAAGCTTGGACACGCGGTTCTGTTCCGGAGGCAAAATTTCCCTAGGAAAGAGACCA 2527  
Db 2325 CAAGCTTGGACACGCGGTTCTGTTCCGGAGGCAAAATTTCCCTAGGAAAGAGACCA 2384  
Qy 2528 GACTTTTCTAATGTGGTCCAAATCGGATCACTGTGAGTGGACTCTAGAAAGCACTGAG 2587  
Db 2385 GACTTTTCTAATGTGGTCCAAATCGGATCACTGTGAGTGGACTCTAGAAAGCACTGAG 2444  
Qy 2588 CTCCCTGTCTGGAAGTATTTAAGAAAGGCTGGGCCAGGACAGTGGCTCAAGCTGT 2647  
Db 2445 CTCCCTGTCTGGAAGTATTTAAGAAAGGCTGGGCCAGGACAGTGGCTCAAGCTGT 2504  
Qy 2648 AATCCACAGATTTGGGAGCGGAGCGAGCTCACTGAGGTGAGGATTTGAGAAACA 2707  
Db 2505 AATCCACAGATTTGGGAGCGGAGCGAGCGAGTCACTGAGTGGAGTGGAGTGGAGAA 2564  
Qy 2708 GCTTGGCAACATGTTGAAACCTCATCTCTACTATAAAATACAAAAATTAGCCAGCGTGG 2767  
Db 2565 GCTTGGCAACATGTTGAAACCTCATCTCTACTATAAAATACAAAAATTAGCCAGCGTGG 2624  
Qy 2768 TGGCAGGTGCTGTAAATCCAGCTACTTGGGAGGCTGAGGATGAGATCACTTAAACCT 2827  
Db 2625 TGGCAGGTGCTGTAAATCCAGCTACTTGGGAGGCTGAGGATGAGATCACTTAAACCT 2684

QY 2828 GAGAGGAGAGGTTACAGTGAGCCAGAGATCGTCCACTGATTCAGCTCCAGCTGGCGGACAGA 2887  
|||||  
Db 2685 GAGAGGAGAGGTTACAGTGAGCCAGAGATCGTCCACTGATTCAGCTCCAGCTGGCGGACAGA 2744  
|||||  
QY 2888 GCAAGACTCTGTCTC 2902  
|||||  
Db 2745 GCAAGACTCTGTCTC 2759  
|||||

## RESULT 11

AB052149  
LOCUS  
DEFINITION  
Macaca fascicularis brain cDNA, clone:QccE-20783, similar to human  
caytaxin (ATCAY), mRNA, NM\_033064.

AB052149  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AB052149.1 GI:11611578  
fis (full insert sequence)  
Macaca fascicularis (crab-eating macaque)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Cercopithecoidea; Cercopithecinae; Macaca.

## REFERENCE

Osada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirata, M.,  
Suto, Y., Hirai, M., Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.  
Assignment of 118 novel cDNAs of cynomolgus monkey brain to human  
chromosomes  
Gene 275 (1), 31-37 (2001)

## JOURNAL

Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.  
Direct Submission  
Submitted (05-DEC-2000) Katsuyuki Hashimoto, National Institute of  
Infectious Diseases, Division of Genetic Resources, 23-1, Toyama

## REFERENCE

1-chome, Shinjuku-ku, Tokyo, 162-8640, Japan  
(E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/,  
Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)

## COMMENT

Lab host: TOP10  
Vector: pME18S-FL3 (Acc.No. AB009864)  
R. Site1: DraIII (CACTGTGTG)  
R. Site2: DraIII (CACCATGTG)  
Description: 1st strand cDNA was primed with an oligo (dT) primer  
using specific 5' and 3' primers and amplified by PCR. The PCR  
product was digested with SfiI and size selection was performed to  
exclude fragments <1.5kb. The SfiI-digested PCR product was cloned  
into distinct DraIII sites of pME18S-FL3. XhoI sites just outside  
the DraIII sites can be used to isolate the cDNA insert. Libraries  
were constructed by Sugano et al. (University of Tokyo, Institute of  
Medical Science). Custom primer used for sequencing  
(5' end primer [CTTCTGCTTAAAGCTGCG];  
3' end primer [CGACCTGCAGCTCGACACA]).

## FEATURES

Location/Qualifiers  
1..2787  
/organism="Macaca fascicularis"  
/mol\_type="mRNA"  
/db\_xref="taxon:9541"  
/clone="QccE-20783"  
/sex="male"  
/tissue="cerebellum cortex"  
/clone\_lib="macaque brain cDNA library QccE"  
/dev\_stage="adult"  
380..1495  
/note="similar to human caytaxin (ATCAY), mRNA, NM\_033064"

## CDS

/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="BAB19004.1"  
/db\_xref="GI:11611578"  
/translation="MGTTEATLRMENVDKSEWDELDPLPRLPSTGVLLGSPVEDT  
SSPPNLFNGAKRKLTVAEDINISLDQSEGLSDPDLDTEDTDLVDNDITETPD  
ETDSELPFGNGELNWDGDDTVPATKNMFGSADLFGDGTDEGANGRLWRITVIG  
EQRHRLDMRPYMKVYTHGGYEGELNIVFAACFLPDSLSLPLHYIMENFLVY  
ISSLELLVAEDYIMVYLNAGTPRRMPGIGWLKKCYQIMDRRLRNKLSLIIVHPSWF

INTVLAI SRPFISVKFINKIQVHSLDEQLIPMEHVQIPDCVLQYBEERLKARRES  
ARPOEFVMPRSEKPEVAPVENRSPVTEQETSMS"

## ORIGIN

Query Match 84.4%; Score 2463; DB 8; Length 2787;  
Best Local Similarity 93.0%; Pred. No. 0;  
Matches 2669; Conservative 0; Mismatches 110; Indels 90; Gaps 5;  
QY 1 GCCGAGCCTCTGCCAGCCCTGAGCTGGGAAGAGCAGCTACCTCCGAGGACGGGCGCGCA 60  
Db 8 GCCGAGCCTCTGCTAGCCCTGAGCTGGGAAGAGCAGCTACCTCCGAGGACGGGCGCGCA 67  
QY 61 GCGGCGCGCGGATGAGAGGGGCGCAGCGCAGCCCGCGCTGGGAGAGCCACCGCTAAC 120  
Db 68 GCGGCGTGGCGATGAGAGCGGCGCAGCGCAGCCCGCGAGGGGAGGCCACCGCTAAC 127  
QY 121 CTTGCACCCACCCACCCCTTGCACAAAGAGCTGGCGGGCGCTGGCCAGCTGCCCTGGG 180  
Db 128 CTTGCACCCACCCACCCCTTGCACAAAGAGCTGGCGGGCGCTGGCCAGCTGCCCTGGG 187  
QY 181 TGACCTTCTCTCGATGCGAATTCGCCCTCGAGCATCTCTTCTCTCTAGGCTCTGAA 240  
Db 188 TGAACCTTCTCTCGATGCGAAGCCCTCTCTGGAGCATCTCTTCTCTCCAGGCTCTGAA 247  
QY 241 GCGCGGGAGCGTGAAGCGATGCCAGCTGCACCCGCGCAGGCTCGCTTTGTTGGCA 300  
Db 248 GGCCTGGGAGCGTGAAGCGATGCCAGCTGCACCCGCGCAGGCTCGCTTTGTTGGCA 307  
QY 301 GTAAGGAGAGGCTGTCTCAGCTGCGAGGGGTTCCTCTGCTTCAAGCCAGTGCCTC 360  
Db 308 GTAAGGAGAGGCTGTCTCAGCTGCGAGGGGTTCACCTCTGCTTCAAGCCAGTGCCTC 367  
QY 361 TTCCAGCTCCCATGGGACCAACCGAGCCAGCTCCCGATGGAAGAGAGCTGAGCTGAG 420  
Db 368 TTCCAGCTCCCATGGGACCAACCGAGCCAGCTCCCGATGGAAGAGAGCTGAGCTGAG 427  
QY 421 CAGGAATGGCAGGACGAAGATCTTCCAGGGCCACTCCCGAGAGAGACGGGGGTGCAACTG 480  
Db 428 GAGGAATGGCAGGACGAAGATCTTCCAGGGCCACTCCCGAGAGAGACGGGGGTGGAATG 487  
QY 481 CTTGGCAGCCCGTGGGAAGACATCTCTCTCTCCCAACACGCTAAATTTCAACGGAGCG 540  
Db 488 CTTGGCAGCCCGTGGGAAGACATCTCTCTCTCCCAACACGCTAAATTTCAACGGAGCG 547  
QY 541 CATGCTAAGAGAGAGAGCTGGTGGCCCGCAGAGATCAATTTCTCTGATCAGAGTGA 600  
Db 548 CATGCTAAGAGAGAGAGCTGGTGGCCCGCAGAGATCAATTTCTCTGATCAGAGTGA 607  
QY 601 GGGTCCCTGCTGTCCGATGACTTCTTGATATACCCCTGATGACCTGGATATTAACCTGGAT 660  
Db 608 GGGTCCCTGCTGTCCGATGACTTCTTGATATACCCCTGATGACCTGGATATTAACCTGGAT 667  
QY 661 GACATCGAGAGACCCCGATGAGACCGAATCTCGCTGGAGTTCTCTGGGAATGGCAACGACTG 720  
Db 668 GACATCGAGAGACCCCGATGAGACTGACTCGCTGGAGTTCTCTGGGAATGGCAACGACTG 727  
QY 721 GAGTGGGAGAGAGACACCCCGTGGCCACCGCCAGAGACATGCCCCGGGAGACGCGCGAT 780  
Db 728 GAGTGGGAGAGAGACACCCCGTGGCCACCGCCAGAGACATGCCCCGGGAGACGCGCGAT 787  
QY 781 CTATTTGGGAGAGAGACACCCCGTGGCCACCGCCAGAGACATGCCCCGGGAGAG 840  
Db 788 CTATTTGGGAGAGAGACACCCCGTGGCCACCGCCAGAGAGAGCGCGGTGGCGGCTGTGGCGGAG 847  
QY 841 GTGATCATCGGGGAGAGAGACCGGTATAGACTTGCACATGATCCGCGCTTACATGAAA 900  
Db 848 GTGATCATCGGGGAGAGAGACCGGTATAGACTTGCACATGATCCGCGCTTACATGAAA 907  
QY 901 GTGCTCACCCACCGAGGGTACTACGGCGAAGCGCTCAACCGCATCATGCTTTCGACGCC 960  
Db 908 GTGCTCACCCACCGAGGGTACTATGGCGAAGCGCTCAACCGCATCATGCTTTCGACGCC 967  
QY 961 TGCTTCTTCCAGAGACGAGCGCTCCCGGACTACCATCATCATCATGAGAACTCTTCTCTG 1020

	Db	968	TGCTTCTTCAGACAGCAGCGCTCCCAGACTACCATCATCATGSGAGAACCCTCTCCCTG	1027
	Qy	1021	TACGTCAATCAGCAGCTTAGAGCTCCTGGTGCTGAGCACTACATGATCGTGTACTCTGAAC	1080
	Db	1028	TACGTCAATCAGCAGCTTAGAGCTCCTGGTGCTGAGCACTACATGATCGTGTACTCTGNAC	1087
	Qy	1081	GGTGCCACGCCCGCGGAGGATGCCTGGAAATCGGCTGGCTGAAGAAAGTGTCTACCAGATG	1140
	Db	1088	GGCGCCACGCCCGCGGAAGGATGCCGGGATCGGCTGGAAGAAAGTGTCTACCAGATG	1147
	Qy	1141	ATCGACCGGAGTTGCGGAAAAAAGTCAAGTCTTTGATCATCGTTCACCCCTGTGTGTTTC	1200
	Db	1148	ATCGATCGGAGATTGCGGAAAAAACCTGAAGTCTTTGATCATCGTTCACCCCTGTGTGTTTC	1207
	Qy	1201	ATTCGGACTGTGTCGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTTCATCAACAGATC	1260
	Db	1208	ATTCGGACTGTGTCGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTTCATCAACAGATC	1267
	Qy	1261	CAGTACGTGCA CAGCTTTGGAGACCTTGGAGCAA CTCTATCCCTATGGAACACGTCCAGATC	1320
	Db	1268	CAGTACGTGCA CAGCTTTGGAGACCTTGGAGCAA CTCTATCCCTATGGAACACGTCCAGATC	1327
	Qy	1321	CCAGACTGCGTCTTGCAATACGAAGAGGAAAAGACTGAAGCCGAGGAGGAGCGCGAGG	1380
	Db	1328	CCAGACTGCGTCTTGCAATACGAAGAGGAAAAGACTGAAGCCGAGGAGGAGCGCGAGG	1387
	Qy	1381	CCCCAGCCGGAGTTTGTGCTGCCAGAGTCTGAAGAGAAAGCCAGAGGTGGCACCAAGTGGAA	1440
	Db	1388	CCCCAGCCGGAGTTTGTGATGCCCAGATCTGAAGGAAGCCAGAGGTGGCACCAAGTGGAA	1447
	Qy	1441	AACAGGTCTGCTCTGGTCTCAGAAAGTACAGAAA CAAGCATGTCTCTCAGGCGCAGTGCAGC	1500
	Db	1448	AACAGGTCTGCTCTGGTCTCAGAAAGTACAGAAA CAAGCATGTCTCTCAGGCGCAGTGCAGC	1507
	Qy	1501	ATAACAAGGACATGGAAAGAGATTCAGAGTGCAGAAAA CCTCTGTTCAGACGCCCACTG	1560
	Db	1508	ATAACAAGGACATGGAAAGAGATTCAGAGTGCAGAAAA CCTCTGTTCAGACGCCCACTG	1567
	Qy	1561	GCCCGAGATCTCATCTCGCTCATCTCTGAGTGCCTTAATCTTCGAAGGTGGCCGCCCTCC	1620
	Db	1568	GCCCGAGATCTCATCTCGCTCATCTCTGAGTGCCTTAATCTTCGAAGGTGGCCGCCCTCC	1627
	Qy	1621	GTTTCATCTCTGAAACCCAGCATCTTTTCAGCTGCTTTGAAAA CATTTATTTTTTTTTT	1680
	Db	1628	GTTTCATCTCTGAAACCCAGCATCTTTTCAGCTGCTTTGAAAA CATTTATTTTTTTTTT	1684
	Qy	1681	TACGATGCA GATATTTGTGGTTTCCGAA AAGGCCCCAGCTCTGAGCCCCCTCACCCCTCC	1740
	Db	1685	TACGATGCA GATATTTGTGGTTTCCGAA AAGGCCCCAGCTCTGAGCCCCCTCACCCCTCC	1744
	Qy	1741	ACATCTCAGAACTCTCAGCCGAGGAAGGCAAGAGCCAGGGGTGCCCGCTGGCGTC	1800
	Db	1745	ACATCTCAGAACTCTCAGCCGAGGAAGGCAAGAGCCAGGGGTGCCCGCTGGCGCTC	1804
	Qy	1801	GGTGGCTCTCGCTCTCGTCCAGCCCTCTGTGGTTCAGAGTGGATACAAGATTCAGACC	1860
	Db	1805	GGTGGCTCTCGCTCTCGTCCAGCCCTCTGTGGTTCAGAGTGGATACAAGATTCAGACC	1864
	Qy	1861	CTTCTCTTGTCTGTGTCACCCGCTCAGGTTTGGAGCCA CAGACCCACCGCCA CCCCGGCT	1920
	Db	1865	CTTCTCTTGTCTGTGTCACCCGCTCTGGGTTTGGAGCCA CAGACCCACCGCCA CCCCGGCT	1924
	Qy	1921	GGGTCTCGCTCTTTCTGTGCTTTTCCCTCCAGAA TGGCGGCTCTCAGACCTAGAGCTCA	1980
	Db	1925	GGGTCTCGCTCTTTCTGTGCTTTTCCCCCAGAACGCGGCTCTGAGACTAGAGCTCA	1984
	Qy	1981	ACCCCCCTATGAGGCCACGCTCTGGGGTAGCTCCTGCACCTCCGACCTTATGTCCAAAT	2040
	Db	1985	ACCCCC- AGGAGGCCACGCTCTGAGGTACCTCTGGCTCTGACCTTACGTCCAAAT	2043
	Qy	2041	TCAACCCATGTTTTTTCATTTGACCGGCCCTTTCTCGTCTCAATATGACACCCAGCTCC	2100

Db	2044	TCACACCGGTGGTTTTTTCATTTCACCGCGCCCTCTCTCGCTCATATAATGACACCCAGCTCC	2103
Qy	2101	TTTGAGAGGATCAGAGCCCAATTCACAGAAGAGCGCTGCCAACCATCTTGTCCTCG	2160
Db	2104	TTTGAGAGGATCAGAGCCCAATTCACAGAAGAGCTGCTGCCAACCATCTTGTCCTCTG	2163
Qy	2161	ATTGCAAAATGACACCCCCAGTAATCTAGAACATTTCTCAAGCCCTTTAACTCAGATGTCA	2220
Db	2164	ATTGCAAAACGACACCCCGTAATCTAGAACATTTCTAAGCCCTTTAACTCAGATGCCA	2223
Qy	2221	AGCCACGGGCAAAACCCCGTCAATACCTCCCAACCAAGGAATGAGATATGTGGACCTCACT	2280
Db	2224	AGCCACAGGCGCAAAACCCAGTCAATACCTCCCAACCAAGGAATGAGATATGTGGACCTCACT	2283
Qy	2281	GCTCCCCCAACCCAGCGTCAAGCTGGGAACGCCAACCGTTGTCGGTGGGAACAGCAG	2340
Db	2284	GCTCCCCCAACCCGGCATCAGGCTGGGAACGCCAACCGTTGTCGGTGGGAACAGCAG	2343
Qy	2341	AGGCTCAGAACTGGCTCTGAAATAGGCAGACTAGCAAGAGGAGATACAGGTTATCGG	2400
Db	2344	AGGCTCAGAAAGTGGCTCTGAAAT-----AGACCTAGCNAAGGAGATACAGGGTAT---	2396
Qy	2401	CGGTTTGAGTGTTCAGAACTCATTCCGGGAAGATAAATCCAGTGCCTGSCCGCAGCCAC	2460
Db	2397	-----	2396
Qy	2461	CTGCATTCAAAGCTTGGACACGCGGGTCTTGTTGCGGAGGCAGAAATTCCTCTAGGAAAAA	2520
Db	2397	-----CCACGCGGTTCTTGTTGCGGAGGCAGAAATTCCTCTGGAAAAA	2438
Qy	2521	GAGACAGACTTTTCTAATGTGTCCAAATGCCGATCACTGTGCAGATGACTCTAGAAG	2580
Db	2439	GAAGACAGACTTTTCTAATGTGTCCAAATGCCGCTCACTGGTCAGATGACTCTAGAAG	2498
Qy	2581	CACGTAGCTCCCTGTCTCTGGAAAGTATTTAAGAAAGGCTGGGCCAGGACGATGCTCA	2640
Db	2499	CAGTGAAGCTCCCTGTCTCTGGAAAGTATTTAAGAAAGGCTTGAAGGCGGACGATGCTCA	2558
Qy	2641	CGCTGTPAATCCCAAG-ACTTTGGAGCGCGAGCGCGGATCACTGAGGTGAGCAGTT	2699
Db	2559	CGCTGTPAATCTTGGACATTTGGAGCGCGAGGTGGCGGATCCCTCTGAAGTCAAGATT	2618
Qy	2700	TGAGAACAGCGCTGGCCAAACATGTGTGAACCTCATCTCTACTAAAAATACAAAAATTAGCC	2759
Db	2619	CGAGACGAGCTGGCCAAACATGTGTGAACCTCATCTCTACTAAAAATACAAAAATTAGCC	2678
Qy	2760	AGGCGTGGTGGCAGGTGCTGTAAATCCCAAGCTTATTGGGAGGCTGAGGCATGAGAAATCAC	2819
Db	2679	AGGCATGGTGGCAGGTGTCTGTAAATCTCAGTACTTTGGGAGGCTAAGGCACGAGAAATCAC	2738
Qy	2820	TTAAACCTGAGAGGCGAGAGGTTACGTAGCCCAAGATCGTGCCACTGCA	2868
Db	2739	TTAAACCTGAGAGGCGAGAGGTTGCAGTGAAGCAGCAAGATCGTGCCACTGCA	2787

RESULT 12  
CQ842483  
LOCUS  
CQ842483 linear PAT 02-AUG-2004  
DEFINITION  
Sequence 1130 from Patent EP1440981.  
ACCESSION  
CQ842483  
VERSION  
CQ842483.1 GI:50894270  
KEYWORDS  
Homo sapiens (human)  
SOURCE  
Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
REFERENCE  
1. Isegai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,  
Yamamoto, J., Isono, Y., Nagai, K. and Irie, R.  
Full-length human cdna  
Patent: EP 1440981-A 1130 28-JUL-2004;  
Research Association for Biotechnology (JP)

FEATURES		Location/Qualifiers	
source	1..2786	/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
ORIGIN			
Query Match		80.8%; Score 2357.6; DB 6; Length 2786;	
Best Local Similarity		95.2%; Pred. No. 0;	
Matches 2492; Conservative		0; Mismatches 4; Indels 122; Gaps 1;	
QY	300	AGTAAGGAGGAGGAGGCTGCTCAGCTGCAGAGGAGGCTCATCCCTGCTTCAAGCCAGTGCCT	359
DB	1	AGTAAGGAGGAGGAGGCTGCTCAGCTGCAGAGGAGGCTCATCCCTGCTTCAAGCCAGTGCCT	60
QY	360	CTTCCAGCTCCCATGGGACACACCGAAGCCAGCTCCGGATGGAAAAAGTGGACGTGAA	419
DB	61	CTTCCAGCTCCCATGGGACACACCGAAGCCAGCTCCGGATGGAAAAAGTGGACGTGAA	120
QY	420	GGAGGAATGGCAGGACGAGATCTTCCAGGACCACTCCAGAGAGACACGGGGTGGAACT	479
DB	121	GGAGGAATGGCAGGACGAGATCTTCCAGGACCACTCCAGAGAGACACGGGGTGGAACT	180
QY	480	GCTTGGCAGCCGGTGGAGACACATCTCTCTCCCAACACGCTAAATTTCAACGGAGC	539
DB	181	GCTTGGCAGCCGGTGGAGACACATCTCTCTCCCAACACGCTAAATTTCAACGGAGC	240
QY	540	GCATCGTAAGAGGAAGACGCTGGTGGCCAGAGATCAACATTTCTCTGGATCAGAGTGA	599
DB	241	GCATCGTAAGAGGAAGACGCTGGTGGCCAGAGATCAACATTTCTCTGGATCAGAGTGA	300
QY	600	GGGGTCCCTGCTGCTCGATGACTTTCTTGATACCCCTGATGACCTGGATATTAACGTGA	659
DB	301	GGGGTCCCTGCTGCTCGATGACTTTCTTGATACCCCTG-----	338
QY	660	TGCATCGAGACCCCGATGAGACCGACTCGCTGGAGTTCCTGGGGAATGGCAACGAAC	719
DB	339	-----	338
QY	720	GGAGTGGGAAGACGACACCCCGTGGCCACCGCAAGAACATGCCGGGACAGCGCGA	779
DB	339	-----ATGCCGGGACAGCGCGA	358
QY	780	TCTATTTGGGACGGACGACGAGGAGACGGCAGCGCGCTGTGGCGGAC	839
DB	359	TCTATTTGGGACGGACGACGAGGAGACGGCAGCGCGCTGTGGCGGAC	418
QY	840	AGTGATCATCGGGAGCAGACGCGTATAGACCTGACATGATCCGGCCCTTAATGAA	899
DB	419	AGTGATCATCGGGAGCAGACGCGTATAGACCTGACATGATCCGGCCCTTAATGAA	478
QY	900	AGTGGTCAACCCAGGAGGTACTACGGGAGAGGCTCAACGCCATCATGCTTCCGACG	959
DB	479	AGTGGTCAACCCAGGAGGTACTACGGGAGAGGCTCAACGCCATCATGCTTCCGACG	538
QY	960	CTGCTTCTTCAGACAGCAGCTCCCGACTACCACTACATCATGAGAACTCTTCT	1019
DB	539	CTGCTTCTTCAGACAGCAGCTCCCGACTACCACTACATCATGAGAACTCTTCT	598
QY	1020	GTACGTATCAGACGTTAGAGCTCTGTGGTGGAGTACATGATCGTGATGCTGAA	1079
DB	599	GTACGTATCAGACGTTAGAGCTCTGTGGTGGAGTACATGATCGTGATGCTGAA	658
QY	1080	CGGTGCCACGCCCGCGGAGATGCTGGATCGGCTGGCTGAGAGATGCTACAGAT	1139
DB	659	CGGTGCCACGCCCGCGGAGATGCTGGATCGGCTGGCTGAGAGATGCTACAGAT	718
QY	1140	GATCGACCGGAGTTGGGAAAAAGCTGAAGTCTTGTATCATGCTCAACCCCTCGTGT	1199
DB	719	GATCGGCGGAGTTGGGAAAAAGCTGAAGTCTTGTATCATGCTCAACCCCTCGTGT	778
QY	1200	CATTGGGACTGTGGGCACTCTCGCCCTTTCATCAGCGCTCAAGTTTCATCAACAGAT	1259
DB			
QY	779	CATTGGGACTGTGGGCACTCTCGCCCTTTCATCAGCGTCAAGTTTCATCAACAGAT	838
DB	1260	CCAGTACGTGCACAGCTTGGAGAGCTGGAGCAACTCATCCCTATGGAACAGTCCAGAT	1319
DB	839	CCAGTACGTGCACAGCTTGGAGAGCTGGAGCAACTCATCCCTATGGAACAGTCCAGAT	898
QY	1320	CCAGACTGCGTCTGCAATACGAAGAGGAAAGACTGAAGGCGAGGAGGAGGCGCGAG	1379
DB	899	CCAGACTGCGTCTGCAATACGAAGAGGAAAGACTGAAGGCGAGGAGGAGGCGCGAG	958
QY	1380	GGCCAGCGGAGTTGTGCTGCCAGAGTCTGGAAGAGCCAGAGGTGGCAACAGTGA	1439
DB	959	GGCCAGCGGAGTTGTGCTGCCAGAGTCTGGAAGAGCCAGAGGTGGCAACAGTGA	1018
QY	1440	AAACAGGTCTGCTGTGCTCAGAGATCAGGAAACAGCAATGCTCTGAGGCGAGTGA	1499
DB	1019	AAACAGGTCTGCTGTGCTCAGAGATCAGGAAACAGCAATGCTCTGAGGCGAGTGA	1078
QY	1500	CATAACAAAGGACATGGAAGAGATTCCAGATGCCAGAAAACTCTGTCAAGCGCCACT	1559
DB	1079	CATAACAAAGGACATGGAAGAGATTCCAGATGCCAGAAAACTCTGTCAAGCGCCACT	1138
QY	1560	GGCCCGAGATCTCATCTGCTCATCTGAGTCCCAATCTTCAAGGGTGCAGCCCTC	1619
DB	1139	GGCCCGAGATCTCATCTGCTCATCTGAGTCCCAATCTTCAAGGGTGCAGCCCTC	1198
QY	1620	GGTTCATCTCTGAAACCCAGCATCTCTTTCAGTCTTGAATAATTTT	1679
DB	1199	GGTTCATCTCTGAAACCCAGCATCTCTTTCAGTCTTGAATAATTTT	1258
QY	1680	TTAAGCATGAGTATTTCTGGTTCAGAAAAAGGCGCAGCTCTCAGCCCTCAGCCCTC	1739
DB	1259	TTAAGCATGAGTATTTCTGGTTCAGAAAAAGGCGCAGCTCTCAGCCCTCAGCCCTC	1318
QY	1740	CACACTCAAGAACTCTCAGCGAGGAGGCAAGAACGCGAGGGTGGCCGCGTGGGT	1799
DB	1319	CACACTCAAGAACTCTCAGCGAGGAGGCAAGAACGCGAGGGTGGCCGCGTGGGT	1378
QY	1800	CGGTGGCTCGCTCTGCTCGAGCCCTTGTGTGAGTGGATACAGATTCAAGAC	1859
DB	1379	CGGTGGCTCGCTCTGCTCGAGCCCTTGTGTGAGTGGATACAGATTCAAGAC	1438
QY	1860	CTTCTCTTGTCTGTACCCGCTCAGGTTGAGGACACAGACACCCACCGCCCGC	1919
DB	1439	CTTCTCTTGTCTGTACCCGCTCAGGTTGAGGACACAGACACCCACCGCCCGC	1498
QY	1920	TGGGTGCGTCTTCTGCTGCTTCCCTCAAGATGCGGCTCAGACCTAGAAGTCT	1979
DB	1499	TGGGTGCGTCTTCTGCTGCTTCCCTCAAGATGCGGCTCAGACCTAGAAGTCT	1558
QY	1980	AACCCCTATGAGGCGCACCTCTGGGTAGTCTGACCTCGAGCTTATGTCAAAT	2039
DB	1559	AACCCCTATGAGGCGCACCTCTGGGTAGTCTGACCTCGAGCTTATGTCAAAT	1618
QY	2040	TTCAACCCATGTTTTCATTTGACCGCCCTTCTCGCTATAATGACACCCAGTCT	2099
DB	1619	TTCAACCCATGTTTTCATTTGACCGCCCTTCTCGCTATAATGACACCCAGTCT	1678
QY	2100	CTTTGAGGAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2159
DB	1679	CTTTGAGGAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1738
QY	2160	GATTGCAAAATGACACCCAGTAACTTAGAACTTCTCAAGCCCTTTAACTCAGATCT	2219
DB	1739	GATTGCAAAATGACACCCAGTAACTTAGAACTTCTCAAGCCCTTTAACTCAGATCT	1798
QY	2220	AAGCCACCGGCAAAACCCGCTCAATCTCCCAAGAGGATGATGTGGAGCTTCAAC	2279
DB	1799	AAGCCACCGGCAAAACCCGCTCAATCTCCCAAGAGGATGATGTGGAGCTTCAAC	1858
QY	2280	TGCTCCCGCAACCCAGGCTCAGGCTGGGACACGCGCAACGCTGTTCGGGTTGGAAACG	2339
DB	1859	TGCTCCCGCAACCCAGGCTCAGGCTGGGACACGCGCAACGCTGTTCGGGTTGGAAACG	1918



QY	2340	GAGGCTCAGAACTGGCTCTGAATAGGACAGCTAGCAAGAGAGAAAGATACAGGGTATCG	2399
Db	1919	GAGGCTCAGAACTGGCTCTGAATAGGACAGCTAGCAAGAGAGAAAGATACAGGGTATCG	1978
QY	2400	GGCGTTTGAGTGTTCAGAAAGTCAATTCGGGAAGATAAATCCAGTGGCTGGCGCAGCCA	2459
Db	1979	GGCGTTTGAGTGTTCAGAAAGTCAATTCGGGAAGATAAATCCAGTGGCTGGCGCAGCCA	2038
QY	2460	CCTGCATTCAAAGCTTGGACACGGGTTCTTGTCGGGAGGCAAAATTTCCCTAGGAAA	2519
Db	2039	CCTGCATTCAAAGCTTGGACACGGGTTCTTGTCGGGAGGCAAAATTTCCCTAGGAAA	2098
QY	2520	AGAAGACAGACTTTTCTAATGTGGTCCAAATGGGATCACTGTGCAGATGGACTCTAGAA	2579
Db	2099	AGAAGACAGACTTTTCTAATGTGGTCCAAATGGGATCACTGTGCAGATGGACTCTAGAA	2158
QY	2580	GCATGAGCTCCCTGTCTCTGGAAGTATTTAAGAAAAGGCTGGGCCAGGCAGATGGCTC	2639
Db	2159	GCATGAGCTCCCTGTCTCTGGAAGTATTTAAGAAAAGGCTGGGCCAGGCAGATGGCTC	2218
QY	2640	ACGCTCTGAATCCAGACTTTGGGAGCGCGAGCGCGATCACTGAGGTGAGGAGTT	2699
Db	2219	ACGCTCTGAATCCAGACTTTGGGAGCGCGAGCGCGATCACTGAGGTGAGGAGTT	2278
QY	2700	TGAGAACAGCTGGCCAACTGATGAACTCTCTACTTAAATAACAAAATTTAGCC	2759
Db	2279	TGAGAACAGCTGGCCAACTGATGAACTCTCTACTTAAATAACAAAATTTAGCC	2338
QY	2760	AGCGGTGGTGGAGGTGCTGTAAATCCAGCTACTTGGGAGGCTGAGGATGAGATCAC	2819
Db	2339	AGCGGTGGTGGAGGTGCTGTAAATCCAGCTACTTGGGAGGCTGAGGATGAGATCAC	2398
QY	2820	TAAACCTGAGCGCAGAGGTTACGTGAGCGCAAGATCGTCCACTGCATTCAGCGCTCG	2879
Db	2399	TAAACCTGAGCGCAGAGGTTACGTGAGCGCAAGATCGTCCACTGCATTCAGCGCTCG	2458
QY	2880	GCAGACAGCAAGACTCTGTCTCAAAAAAATAAAAAA	2917
Db	2459	GCAGACAGCAAGACTCTGTCTCAAAAAAATAAAAAA	2496
RESULT 13			
AKI25457			
LOCUS			
DEFINITION	AKI25457 2786 bp mRNA linear PRI 09-SEP-2003		
	Homo sapiens cDNA FLJ43468 fis. clone OCBFP2037068, moderately		
	similar to BCL2/adenovirus E1B 19-kDa protein-interacting protein		
	2.		
ACCESSION	AKI25457		
VERSION	AKI25457.1 GI:34531558		
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Oshima, A., Takahashi-Pujji, A., Tanase, T., Imose, N., Takeuchi, K.,		
	Arita, M., Muraahino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R.,		
	Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J.,		
	Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K.,		
	Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H.,		
	Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A.,		
	Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K.		
	and Isogai, T.		
	NEDO human cDNA sequencing project		
	Unpublished		
JOURNAL	2 (bases 1 to 2786)		
REFERENCE	Isogai, T. and Yamamoto, J.		
AUTHORS	Direct Submission		
TITLE	Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7		
JOURNAL	Kazusa-Kamataru, Kisarazu, Chiba 232-0818, Japan		
	(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)		

COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.		
FEATURES	Location/Qualifiers		
source	1..2786		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="OCBFP2037068"		
	/tissue_type="brain"		
	/clone_lib="OCBFP2"		
	/dev_stage="fetal"		
	/note="cloning vector: pME18SFL3"		
ORIGIN			
	Query Match	80.8%	Score 2357.6; DB 8; Length 2786;
	Best Local Similarity	95.2%	Pred. No. 0;
	Matches 2492; Conservative	0; Mismatches	4; Indels 122; Gaps 1;
QY	300	AGTAAGGAGGAGGCTGTCTCAGCTGCAGAGGGGTCTATCCCTGCTTCAAGCCAGTGCCT	359
Db	1	AGTAAGGAGGAGGCTGTCTCAGCTGCAGAGGGGTCTATCCCTGCTTCAAGCCAGTGCCT	60
QY	360	CTTCCAGCTCCATGGGGACCCAGAAAGCCACGCTCCGGATGGAAGAAACGTGGAAGTGA	419
Db	61	CTTCCAGCTCCATGGGGACCCAGAAAGCCACGCTCCGGATGGAAGAAACGTGGAAGTGA	120
QY	420	GGAGAAATGGGAGGAGGAGATCTTCCAGGCACTCCAGAGAGAGCGGGGTGGAAC	479
Db	121	GGAGAAATGGGAGGAGGAGATCTTCCAGGCACTCCAGAGAGAGCGGGGTGGAAC	180
QY	480	GCTTGGCAGCCGGTGGAGAGACATCTCTCTCCCAACACGCTTAAATTTCAACGGAGC	539
Db	181	GCTTGGCAGCCGGTGGAGAGACATCTCTCTCCCAACACGCTTAAATTTCAACGGAGC	240
QY	540	GCATCTAAGAGAGAGAGCGTGGTGGCCCGCAGAGATCAACATTTCTCTGGATCAGAGTGA	599
Db	241	GCATCTAAGAGAGAGAGCGTGGTGGCCCGCAGAGATCAACATTTCTCTGGATCAGAGTGA	300
QY	600	GGGGTCCCTGCTGTCGGATGACTTCTTGATACCCCTGATGACCTGATATTAACTGGA	659
Db	301	GGGGTCCCTGCTGTCGGATGACTTCTTGATACCCCTG-----	338
QY	660	TGACATCGAGACCCCGGATGAGACCGGACTCGTGGAGTTCTTGGGGAAATGGCAACGAACT	719
Db	339	-----	338
QY	720	GGAGTGGAGAGAGACACCCCGTGGCCACCGCCAAAGAACATGCCCGGGGAGCAGCGGGA	779
Db	339	-----ATGCCCGGGGAGCAGCGGGA	358
QY	780	TCTATTTGGGAGCGCAGCAGCGAGAGCGCGCCCAACCGCGCGCTGTGTGCGGAC	839
Db	359	TCTATTTGGGAGCGCAGCAGCGAGAGCGCGCCCAACCGCGCGCTGTGTGCGGAC	418
QY	840	AGTGTATCTCGGGGAGCAGAGACCGTATAGACCTGACATGATCCGGCCCTTACATGAA	899
Db	419	AGTGTATCTCGGGGAGCAGAGACCGTATAGACCTGACATGATCCGGCCCTTACATGAA	478
QY	900	AGTGTATCTCGGGGAGGTTACTAGCGGAGAGCGCTCAACGCCATCATCTTCTGCGAGC	959
Db	479	AGTGTATCTCGGGGAGGTTACTAGCGGAGAGCGCTCAACGCCATCATCTTCTGCGAGC	538
QY	960	CTGCTTCTTCCAGACAGCAGCGCTCCCGAGCTTACCACCTACATCATGAGAGAACCTTCTCT	1019
Db	539	CTGCTTCTTCCAGACAGCAGCGCTCCCGAGCTTACCACCTACATCATGAGAGAACCTTCTCT	598
QY	1020	GTACGTCTATCAGCAGCTTAGAGCTCTCTGGTGGCTGAGGACTACATGATCTGATGAGTGA	1079



Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Majek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Murny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fanej, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Grimwood, J.W., Green, E.D., Dickson, M.C., Rodrigues, A.C., Trimmwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2276)

Strausberg, R.

Direct Submission

Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Oct 28, 2003 this sequence version replaced gi:14250564.

Contact: MGC help desk

Email: [cgapsb-r@mail.nih.gov](mailto:cgapsb-r@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;

Web site: <http://www.nisg.nih.gov/>

Contact: [nisc\\_mgc@nih.gov](mailto:nisc_mgc@nih.gov)

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-I., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgoun, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggin, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 5 Row: 0 Column: 7  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

```

FEATURES
source
analysis. Location/Qualifiers
1..2276
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:3163511"
/tissue_type="Brain, neu
/clone_lib="NIH MGC 19"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"

```

ORIGIN						/notes=-vector; poib?/"
	Query Match	62.2%	Score 1815, 8;	DB 8;	Length 2276;	
	Best Local Similarity	96.6%	Pred. No. 0;			
	Matches 1902;	Conservative	0;	Mismatches	2;	Indels 65; Gaps 2;
<hr/>						
Qy	1013	TCTTCCTGTACGTCATCAGCAGCTTACAGACTCCTGGTGCGCTGAGGACTACATCATCTCGTGT				1072
Dd	1	TCTTCCTGTACGTCATCAGCAGCTTACAGACTCCTGGTGCGCTGAGGACTACATCATCTCGTGT				60
<hr/>						
Qy	1073	ACTCGAACGGTGCCACGCCGCCGGGAGGATGCCTCGAATCGGCTGGCTCGAAGAAGTGCT				1132

Db	61	 ACCTGAACGGTGCACGCCCCGCGGAGATGCCTGGAATCGCTGGCTGAAGAGTGCT	120
Qy	1133	 ACCAGATGATCGACCGGAGGTTGCGGAAAAA CCTGAAGTCCTTGATCATCGTCCACCCCT	1192
Db	121	 ACCAGATGATCGACCGGAGGTTGCGGAAAAA CCTGAAGTCCTTGATCATCGTCCACCCCT	180
Qy	1193	 CGTGGTTCAATCGGACTGTGCTGGCCATCTCTCGCCCTTT CATCAGCGTCAAGTTTCATCA	1252
Db	181	 CGTGGTTCAATCGGACTGTGCTGGCCATCTCTCGCCCTTT CATCAGCGTCAAGTTTCATCA	240
Qy	1253	 ACAAGATCCAGTACGTCGCA CAGCTTGGAAAGACCTGGAGCAACTCATCTCTATGGAACA CG	1312
Db	241	 ACAAGATCCAGTACGTCGCA CAGCTTGGAAAGACCTGGAGCAACTCATCTCTATGGAACA CG	300
Qy	1313	 TCCAGATCCAGACTCGGCTCTGCAATACGAAGAGGAAGACCTGAAGGCCAGGAGGGAGA	1372
Db	301	 TCCAGATCCAGACTCGGCTCTGCAATACGAAGAGGAAGACCTGAAGGCCAGGAGGGAGA	360
Qy	1373	 GCGCAGGCCCCAGCCGAGGTTGTGCTCCGCCAGGTCTCAAGAGAAGCCAGAGGTGGCAC	1432
Db	361	 GCGCAGGCCCCAGCCGAGGTTGTGCTCCGCCAGGTCTCAAGAGAAGCCAGAGGTGGCAC	420
Qy	1433	 CAGTGGAAAAACAGGTCTGCTCTGGTCTCAGAAAGATCAGGAAACAAG-	1478
Db	421	 CAGTGGAAAAACAGGTCTGCTCTGGTCTCAGAAAGATCAGGAAACAAGCCTCCAGCCCTGC	480
Qy	1479	----- -----CATGTCCTGA 1488	
Db	481	 ACAGATCAAGCTGATGCCA CGTGCCCCAAGTTCCAACCTTAAGTCACGTTCATGTCCTGA	540
Qy	1489	 GGCGAGCTGAGCATAAACAAGGACATGGAAGAAGATTCCAGATGCCAGAAAACTCTGTC	1548
Db	541	 GGCGAGCTGAGCATAAACAAGGACATGGAAGAAGATTCCAGATGCCAGAAAACTCTGTC	600
Qy	1549	 AGAAGCCCACTGGCCCCAGATCTCATCTCTGCTCATCTCTGAGTCCCAATCTTCCAAAGGT	1608
Db	601	 AGAAGCCCACTGGCCCCAGATCTCATCTCTGCTCATCTCTGAGTCCCAATCTTCCAAAGGT	660
Qy	1609	 GCCAGCCCTCGGTTTCATCTCTGAAACCAGCATCCTTTT CAGCTGCTTGAAACAATTGT	1668
Db	661	 GCCAGCCCTCGGTTTCATCTCTGAAACCAGCATCCTTTT CAGCTGCTTGAAACAATTGT	720
Qy	1669	 ATTTTTTTTTTTTAAAGATGACGATATTGTGCGTTCCAGAAAGGCCCCAGCTCTGAGCC	1728
Db	721	 A-TTTTTTTTTTTTAAAGATGACGATATTGTGCGTTCCAGAAAGGCCCCAGCTCTGAGCC	779
Qy	1729	 CCTCACCTTCCACATCA CGAACTCTCAGCCGAGGAAGCGCAGGGGGTGGC	1788
Db	780	 CCTCACCTTCCACATCA CGAACTCTCAGCCGAGGAAGCGCAGGGGGTGGC	839
Qy	1789	 CGCGTGGCGTGGTGGCCCTCGCTCTGCTCGAGCCCTGTGTGTGTCAGAGCTGGATACA	1848
Db	840	 CGCGTGGCGTGGTGGCCCTCGCTCTGCTCGAGCCCTGTGTGTGTCAGAGCTGGATACA	899
Qy	1849	 AGATTCAAGACCTTCTCTTGTGTGTCACCGCTCCAGGTCGAGCTCGAGCACACCCACC	1908
Db	900	 AGATTCAAGACCTTCTCTTGTGTGTCACCGCTCCAGGTCGAGCTCGAGCACACCCACC	959
Qy	1909	 GCCACCCCGGCTGGGCTCGCTCTTGTGTCCTGTGTCCTTCCAGAAATCGCGCTCAGA	1968
Db	960	 GCCACCCCGGCTGGGCTCGCTCTTGTGTCCTTGTGTCCTTCCAGAAATCGCGCTCAGA	1019
Qy	1969	 CCTAGAGCTCAACCCCCCTATGAGGGCCACGTCCTGGGGTAGCTCTGACCTCGACCT	2028
Db	1020	 CCTAGAGCTCAACCCCCCTATGAGGGCCACGTCCTGGGGTAGCTCTGACCTCGACCT	1079
Qy	2029	 TATGTCAAATTTCA CACCATGGTTTTTTCATTTGACCCGCCCTTCTCGCTCAATG	2088
Db	1080	 TATGTCAAATTTCA CACCATGGTTTTTTCATTTGACCCGCCCTTCTCGCTCAATG	1139
Qy	2089	 ACACCCAGCTCTCTTTTGAGAGGATCAGAGCCCATGTCACAAGAGAGCCGCTGCCAACCAT	2148



Db 1140 ACACCCAGCTCCTTTGAGAGATCAGAGCCATTGTCACAGAGAGAGCCGTCGCCAACCAT 1199  
QY 2149 CTTTGCTCCTCGATTGCAAAATGACACCCAGTAATCTAGAAATCTCAAGCCCTTTA 2208  
Db 1200 CTTTGCTCCTCGATTGCAAAATGACACCCAGTAATCTAGAAATCTCAAGCCCTTTA 1259  
QY 2209 ACTCAGATGTCAGACCCAGGGCAAAACCCGCTCAATCTCCCAACCAAGATGAGATAT 2268  
Db 1260 ACTCAGATGTCAGACCCAGGGCAAAACCCGCTCAATCTCCCAACCAAGATGAGATAT 1319  
QY 2269 GTGAGCTCACTGCTCCTCCCAACCCAGCGTCAGCGTGGGACCGCCAAACGCTGTTCCGGG 2328  
Db 1320 GTGAGCTCACTGCTCCTCCCAACCCAGCGTCAGCGTGGGACCGCCAAACGCTGTTCCGGG 1379  
QY 2329 TTGGAACAGCAGAGGCTCGAAACTGCTGCTGAAATAGGCAGACCTAGCAAGAGGAAGAT 2388  
Db 1380 TTGGAACAGCAGAGGCTCGAAACTGCTGCTGAAATAGGCAGACCTAGCAAGAGGAAGAT 1439  
QY 2389 ACAGGATATCGGCGTTTGAAGTCTTTCAGAGTTCATTCGGGAAGATAAATCCAGTGCCT 2448  
Db 1440 ACAGGATATCGGCGTTTGAAGTCTTTCAGAGTTCATTCGGGAAGATAAATCCAGTGCCT 1499  
QY 2449 GCGCGCAGCACCCTGCAATCAAGCTTGGACAGCGGTTCTGTTGGGAGGCAAAATTT 2508  
Db 1500 GCGCGCAGCACCCTGCAATCAAGCTTGGACAGCGGTTCTGTTGGGAGGCAAAATTT 1559  
QY 2509 CCTAGGAAAAAGAACAGACACTTTTCTAATGTGTCCTAATGCGGATCACTGGTCAGAT 2568  
Db 1560 CCTAGGAAAAAGAACAGACACTTTTCTAATGTGTCCTAATGCGGATCACTGGTCAGAT 1619  
QY 2569 GGAATCTAGAGCACTGAGTCTCTCTCTGGAAGTATTAAAGAAAGGCTGGGCGAG 2628  
Db 1620 GGAATCTAGAGCACTGAGTCTCTCTCTGGAAGTATTAAAGAAAGGCTGGGCGAG 1679  
QY 2629 CAGATGGCTCAGCTGTAATCCAGACTTTGGAGGCGGAGGCGGATCACTGCA 2688  
Db 1680 CAGATGGCTCAGCTGTAATCCAGACTTTGGAGGCGGAGGCGGATCACTGCA 1739  
QY 2689 GGTGAGGAGTTTGAGAACAGCTGGCCAAACATGTTGAAACCTCATCTACTATAAAATAC 2748  
Db 1740 GGTGAGGAGTTTGAGAACAGCTGGCCAAACATGTTGAAACCTCATCTACTATAAAATAC 1799  
QY 2749 AAAAAATTAGCAGGCTGGTGAGGCTGCTGTAATCCAGCTACTTTGGAGGCTGAGGC 2808  
Db 1800 AAAAAATTAGCAGGCTGGTGAGGCTGCTGTAATCCAGCTACTTTGGAGGCTGAGGC 1859  
QY 2809 ATGAGATCACTTAACCTGAGAGGAGGCTTACAGTGGCCAGATGCTGCCACTGCA 2868  
Db 1860 ATGAGATCACTTAACCTGAGAGGAGGCTTACAGTGGCCAGATGCTGCCACTGCA 1919  
QY 2869 TTCCAGCTGGGCGCAGAGCAAGACTCTGCTCAAAAAAATAAAAAA 2917  
Db 1920 TTCCAGCTGGGCGCAGAGCAAGACTCTGCTCAAAAAAATAAAAAA 1968

RESULT 15  
AC011488  
LOCUS AC011488 134308 bp DNA linear PRI 01-AUG-2002  
DEFINITION Homo sapiens chromosome 19 clone CTB-171N13, complete sequence.  
ACCESSION AC011488  
VERSION AC011488.7 GI:22038500  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 134308)  
REFERENCE  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 134308)  
AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission  
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 3 (bases 1 to 134308)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
JOURNAL Direct Submission  
TITLE Submitted (01-AUG-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
COMMENT On Aug 1, 2002 this sequence version replaced gi:9256301.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 100% of Sequence;  
Estimated Total Number of Errors is 0.  
FEATURES  
source  
1..134308  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTB-171N13"  
ORIGIN  
Query Match 49.3%; Score 1438.2; DB 8; Length 134308;  
Best Local Similarity 99.8%; Pred. No. 2.8e-272;  
Matches 1440; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1475 CAAGCATGCTCTGAGGCGACGTGAGCATAAACAAAGGACATGGAAGAATTCAGATGCC 1534  
Db 74150 CTAGCATGCTCTGAGGCGACGTGAGCATAAACAAAGGACATGGAAGAATTCAGATGCC 74209  
QY 1535 AGAAACCTCTGTCAGAGCCCACTGGCCCGCAGATCTCATCTGCTCATCTCTGAGTCCC 1594  
Db 74210 AGAAACCTCTGTCAGAGCCCACTGGCCCGCAGATCTCATCTGCTCATCTCTGAGTCCC 74269  
QY 1595 AATCTTCAAGGGTCCAGCCCTCCGTTTCATCTCTGAAACCCAGCATCTTTTTCAGCTG 1654  
Db 74270 AATCTTCAAGGGTCCAGCCCTCCGTTTCATCTCTGAAACCCAGCATCTTTTTCAGCTG 74329  
QY 1655 CTGAAAAACATGTATTTTTTTTTTAAACGATGAGTATTTGCTGCTTCCAGAAAAAGG 1714  
Db 74330 CTGAAAAACATGTATTTTTTTTTTAAACGATGAGTATTTGCTGCTTCCAGAAAAAGG 74389  
QY 1715 CCCAGCTCTGAGCCCTCACTCCCTTCCACATCTCAGCGGAGGAAGGCAAGAA 1774  
Db 74390 CCCAGCTCTGAGCCCTCACTCCCTTCCACATCTCAGCGGAGGAAGGCAAGAA 74449  
QY 1775 GCGCAGGGGGTGGCCCGGTGGCGTGGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 1834  
Db 74450 GCGCAGGGGGTGGCCCGGTGGCGTGGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 74509  
QY 1835 CAGAGCTGGATACAGATTTCAAGACCTTCTCTTCTTGTGTCAACCCGCTCCAGGTTGGAGC 1894  
Db 74510 CAGAGCTGGATACAGATTTCAAGACCTTCTCTTCTTGTGTCAACCCGCTCCAGGTTGGAGC 74569  
QY 1895 CACAGACACCAACCGCCACCCCGGTGGCTCTGCTCTCTCTCTGCTGCTTCCCTCCAG 1954  
Db 74570 CACAGACACCAACCGCCACCCCGGTGGCTCTGCTCTCTCTCTGCTGCTTCCCTCCAG 74629  
QY 1955 AATCGGCTCAGACCTAGAAAGCTCAACCCCTTATAGGGCCACGCTCTGGGGTAGCTC 2014  
Db 74630 AATCGGCTCAGACCTAGAAAGCTCAACCCCTTATAGGGCCACGCTCTGGGGTAGCTC 74689  
QY 2015 CTGACCTCCGACCTTATGTCCAAATTTCAACCCATGTTTTTCAATTTGACCCGCCCT 2074  
Db 74690 CTGACCTCCGACCTTATGTCCAAATTTCAACCCATGTTTTTCAATTTGACCCGCCCT 74749  
QY 2075 TCTGCTCATATGACACCCAGCTCTCTTTGAGAGATCAGAGCCCATGTCACAAAGAG 2134  
Db 74750 TCTGCTCATATGACACCCAGCTCTCTTTGAGAGATCAGAGCCCATGTCACAAAGAG 74809  
QY 2135 CCGCTGCCAACCATCTCTGCTCCGATTTGCAAAATGACACCCCGAGTATCTAGAACATT 2194

```
|||||
Db 74810 CCCTGCAACCATCTCTGTCTCCGATTGCAAAATGACACCCAGTAATCTAGAACATT 74869
QY 2195 CTCAGCCCTTTAACTCAGATGTCAGGCCCAAGCCGCGCAAAACCCCGTCAATACCTCCACC 2254
Db 74870 CTCAGCCCTTTAACTCAGATGTCAGGCCCAAGCCGCGCAAAACCCCGTCAATACCTCCACC 74929
QY 2255 AAGGAATGAGATATGTGGACCTCACTGCTCCCCCAACCCAGCGTCAGGCTGGGACACGCC 2314
Db 74930 AAGGAATGAGATATGTGGACCTCACTGCTCCCCCAACCCAGCGTCAGGCTGGGACACGCC 74989
QY 2315 AAGCTGTTCCGGTTTGGAAACAGACAGAGGCTCAGAACTGGCTCTGAAATAGGACACCT 2374
Db 74990 AAGCTGTTCCGGTTTGGAAACAGACAGAGGCTCAGAACTGGCTCTGAAATAGGACACCT 75049
QY 2375 AGCAAGAGGAAGATACAGGGTATCGGGCGTTTCAGTGTTCAGAACTCAATTCGGGAAGAT 2434
Db 75050 AGCAAGAGGAAGATACAGGGTATCGGGCGTTTCAGTGTTCAGAACTCAATTCGGGAAGAT 75109
QY 2435 AAATCCAGTGCCTGGCCGACGACCTGTCATTCAAAGCTTGGACCAAGCGGTTCTTGT 2494
Db 75110 AAATCCAGTGCCTGGCCGACGACCTGTCATTCAAAGCTTGGACCAAGCGGTTCTTGT 75169
QY 2495 CGGGAGGCAAAATTCCTAGGAAAGAAAGACAGACTTTTCTAAATGTGTCCAAATGCGG 2554
Db 75170 CGGGAGGCAAAATTCCTAGGAAAGAAAGACAGACTTTTCTAAATGTGTCCAAATGCGG 75229
QY 2555 ATCACTGGTCCAGATGGACTCTAGAAGCACTGAGCTCCCTGTCTCTGGAAGTATTTAAGAA 2614
Db 75230 ATCACTGGTCCAGATGGACTCTAGAAGCACTGAGCTCCCTGTCTCTGGAAGTATTTAAGAA 75289
QY 2615 AAGGCTGGGCCAGGCACGATGGCTCAGCGCTGTAAATCCAGACTTTGGGAGGCCGAGGCA 2674
Db 75290 AAGGCTGGGCCAGGCACGATGGCTCAGCGCTGTAAATCCAGACTTTGGGAGGCCGAGGCA 75349
QY 2675 GCGGATCACTGAGGTGAGGATTTTGAGAAACAGCTGGCCCAACATGGTGAAACCTCATC 2734
Db 75350 GCGGATCACTGAGGTGAGGATTTTGAGAAACAGCTGGCCCAACATGGTGAAACCTCATC 75409
QY 2735 TCTACTAAAAATACAAAAATTAGCCAGGCGTGTGGCAGGTGCTGTATCCGCTACT 2794
Db 75410 TCTACTAAAAATACAAAAATTAGCCAGGCGTGTGGCAGGTGCTGTATCCGCTACT 75469
QY 2795 TGGAGGCTGAGGCATGAGAAATCACTTAAACCTGAGAGGCAGAGGTTACAGTGAGCCAAG 2854
Db 75470 TGGAGGCTGAGGCATGAGAAATCACTTAAACCTGAGAGGCAGAGGTTACAGTGAGCCAAG 75529
QY 2855 ATCGTCCACTGCATTTCCAGCTTGGGCGACAGAGCAAGACTCTGTCTCAAAAAAATAA 2914
Db 75530 ATCGTCCACTGCATTTCCAGCTTGGGCGACAGAGCAAGACTCTGTCTCAAAAAAATAA 75589
QY 2915 AAA 2917
Db 75590 AAA 75592
```

Search completed: March 5, 2006, 07:27:41  
Job time : 14453 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 02:57:19 ; Search time 1671 Seconds  
(without alignments)

11638.284 Million cell updates/sec

Title: US-10-699-941-3

Perfect score: 2918

Sequence: 1 gccgagcctcgcagccct.....tctcaaaaaaaaaaaaaa 2918

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_21:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2001bs:\*
- 6: Geneseq2002s:\*
- 7: Geneseq2002bs:\*
- 8: Geneseq2003s:\*
- 9: Geneseq2003bs:\*
- 10: Geneseq2003cs:\*
- 11: Geneseq2003ds:\*
- 12: Geneseq2004s:\*
- 13: Geneseq2004bs:\*
- 14: Geneseq2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2918	100.0	2918	13	ADQ16276 Human Cay
2	2916.4	99.9	2918	13	ADQ16281 Human Cay
3	2913.8	99.9	4978	12	ADQ85643 Human tum
4	2913.8	99.9	4978	13	ADQ40413 Human tum
5	2896.6	99.3	4976	6	ADQ83973 Human gen
6	2887.6	99.0	3065	3	ADQ00128 Human NIP
7	2870.2	98.4	2894	4	ADQ16751 Human cdn
8	2598.2	89.0	2759	10	ADQ62843 Human cdn
9	2357.6	80.8	2786	12	ADQ63969 Novel hum
10	1438.2	49.3	47219	13	ADQ16283 Human Cay
11	1438.2	49.3	59884	13	ADQ16284 Human Cay
12	1325.8	45.4	2625	10	ADQ62832 Human cdn
13	854.6	29.3	2166	13	ADQ16274 Mouse jit
14	745.2	25.5	790	4	ADQ08588 Human cdn
15	675	23.1	703	5	ADQ80323 DNA encod
16	579	19.8	579	12	ADQ92257 Human aut
17	470	16.1	560	4	ADQ13509 Human cdn
18	420.4	14.4	485	9	ADQ15328 Human adu
19	331.6	11.4	384	2	ADQ59527 Human bra

20	330.4	11.3	5656	12	ADQ21252	Adq21252 Human sof
21	327.6	11.2	5668	13	ADQ82885	Adq82885 Human KIA
22	327.6	11.2	11690	9	ACC45236	Acc45236 Human BMC
23	318.2	10.9	5654	6	ABL62307	Abi62307 Colon ade
24	318.2	10.9	5654	6	ABL68862	Abi68862 Kidney ca
25	318.2	10.9	5654	6	ABL68482	Abi68482 Kidney ca
26	315.8	10.8	2347	4	AAI98076	Aai98076 Human neu
27	302.2	10.4	4235	3	AAQ00129	Aaq00129 Human NIP
28	302.2	10.4	4378	5	ABV27987	Abv27987 Human pro
29	302.2	10.4	4378	5	ABV22147	Abv22147 Human pro
30	259.2	8.9	273	5	ABA13564	Abal3564 Human ner
31	245	8.4	2417	13	ACN40083	Actn40083 Tumour-as
32	243.6	8.3	660	5	ADL39125	Adl39125 Human ova
33	243.4	8.3	2544	4	AAH14059	Aah14059 Human cdn
34	242.2	8.3	139308	8	ADB12769	Adb12769 Human PRK
35	241.8	8.3	96593	11	ADL27146	Adl27146 Human gen
36	241.8	8.3	96595	9	ADA03068	Ada03068 Human PPP
37	241.8	8.3	96595	10	ADB72806	Adb72806 Human PPP
38	241.8	8.3	96596	9	ADA66352	Ada66352 Human PPP
39	241.6	8.3	110000	10	ADE11692	Adel11692 Human can
40	240.8	8.3	138627	12	ADQ97183	Adq97183 Human can
41	240	8.2	100608	13	ABD32769	Abd32769 Human can
42	239.6	8.2	20001	13	ADT77129	Adt77129 Type II d
43	239.6	8.2	32249	4	AAI05336	Aai05336 Human rep
44	239.6	8.2	32249	4	ABL98205	Abi98205 Human tes
45	239.2	8.2	8939	5	ABA18915	Abal8915 Human ner

## ALIGNMENTS

## RESULT 1

ADQ16276

ID ADR16276 standard; cDNA; 2918 BP.

XX AC ADR16276;

XX 21-OCT-2004 (first entry)

XX Human Cayman ataxia cDNA.

XX Human; jittery; ss; gene; Cayman ataxia; ATCAY; KIAA1872;

KW Chromosome 19p13.3; ataxia; myoclonus; dystonia; epileps; myastagmus;

KW SNP; single nucleotide polymorphism.

XX OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 373..1488

FT /\*tag= a

FT /product= "Cayman ataxia protein"

FT replace(1275,6)

FT /\*tag= b

FT /standard\_name= "Single nucleotide polymorphism"

XX US2004146900-A1.

XX 29-JUL-2004.

XX 03-NOV-2003; 2003US-00699941.

XX 01-NOV-2002; 2002US-0422971P.

XX 08-NOV-2002; 2002US-0424973P.

XX (UNMI ) UNIV MICHIGAN.

XX Burmeister M;

XX WPI; 2004-552665/53.

XX P-PSDB; ADR16277.

XX Detecting variant Cayman ataxia polypeptide or nucleic acid sequence in

XX subject, by detecting presence or absence of variant Cayman ataxia

XX

XX

XX

XX

XX

XX

XX

XX

XX

PT polypeptide or nucleic acid in biological sample.

PS Claim 4; SEQ ID NO 3; 98pp; English.

XX The invention relates to detecting variant Cayman ataxia polypeptide or  
XX nucleic acid sequence (ATCAY) in a subject, involving providing a  
XX biological sample from a subject, where the biological sample comprises a  
XX Cayman ataxia polypeptide or nucleic acid, and detecting the presence or  
XX absence of a variant Cayman ataxia polypeptide or nucleic acid in the  
XX biological sample. Also included is a kit comprising a reagent for  
XX detecting the presence or absence of a variant Cayman ataxia nucleic acid  
XX or polypeptide in a biological sample. The human ATCAY gene (previously  
XX isolated as cDNA KIA1872, located on chromosome 19p13.3) was identified  
XX and mapped using the fact that it is the homologue of the mouse Jittery  
XX gene, located on mouse chromosome 10. The method is useful for detecting  
XX variant Cayman ataxia polypeptide or nucleic acid sequence in a subject.  
XX The presence of the variant Cayman ataxia polypeptide or nucleic acid is  
XX indicative of a disorder chosen from ataxia, myoclonus, dystonia,  
XX epilepsy, and nystagmus in the subject. The biological sample is chosen  
XX from blood sample, a tissue sample, urine sample, saliva sample, and an  
XX amniotic fluid sample. The subject is chosen from embryo, foetus, newborn  
XX animal, young animal, and an adult animal. The animal is a human. The  
XX human is an adult female of child-bearing age. The present sequence is  
XX the cDNA encoding the Human ATCAY protein, the allele encoding the  
XX normal, non-disease causing protein.

XX SQ Sequence 2918 BP; 699 A; 868 C; 797 G; 554 T; 0 U; 0 Other;

Query Match 100.0%; Score 2918; DB 13; Length 2918;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2918; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGAGCCTCTGCGAGCCTCTGAGCTGGGAAAGAGAGCTACCTCGGAGGACGGCGCGCA 60  
DB 1 GCGGAGCCTCTGCGAGCCTCTGAGCTGGGAAAGAGAGCTACCTCGGAGGACGGCGCGCA 60  
QY 61 GCGGCGCGCGATGAGAGGGGGCGCAGCGCAGCCCGCGCTGGGGAGCCCAAC 120  
DB 61 GCGGCGCGCGATGAGAGGGGGCGCAGCGCAGCCCGCGCTGGGGAGCCCAAC 120  
QY 121 CCTGCAACCCACCCACCTGCAAAAGAGCTGGCGGGCGCTGGCGAGCTGCTTCTTCTAGGCTCTGAA 240  
DB 121 CCTGCAACCCACCCACCTGCAAAAGAGCTGGCGGGCGCTGGCGAGCTGCTTCTTCTAGGCTCTGAA 240  
QY 181 TGACCTTCTCGGATGCAAGATCCGCGCTGCGAGCATCTCTTCTCTAGGCTCTGAA 240  
DB 181 TGACCTTCTCGGATGCAAGATCCGCGCTGCGAGCATCTCTTCTCTAGGCTCTGAA 240  
QY 241 GCGCGGGAGCGTGAGCGATGCGACAGTCCGCGCGCGAGCGGCTGCGCTTTTGTGCA 300  
DB 241 GCGCGGGAGCGTGAGCGATGCGACAGTCCGCGCGCGAGCGGCTGCGCTTTTGTGCA 300  
QY 301 GTAAGGAGAGAGGCTCTCTCAGCTGCAAGGGGTCTCCCTGCTTCAAGCCAGTGCCTC 360  
DB 301 GTAAGGAGAGAGGCTCTCTCAGCTGCAAGGGGTCTCCCTGCTTCAAGCCAGTGCCTC 360  
QY 361 TTCCAGCTCCCATGCGGACCCAGGACCGCTCCGATGGAACCGTGACGTGAAG 420  
DB 361 TTCCAGCTCCCATGCGGACCCAGGACCGCTCCGATGGAACCGTGACGTGAAG 420  
QY 421 GAGGAATGGCAGGACGAAATTTTCCAGGCCACTCCCAAGAGAGACGGGGTGGAACTG 480  
DB 421 GAGGAATGGCAGGACGAAATTTTCCAGGCCACTCCCAAGAGAGACGGGGTGGAACTG 480  
QY 481 CTTGGCAGCCCGGTGGAAGACATCTCTCTCCCAACACCGTAAATTTCAACGGAGCG 540  
DB 481 CTTGGCAGCCCGGTGGAAGACATCTCTCTCCCAACACCGTAAATTTCAACGGAGCG 540  
QY 541 CATCGTAAGAGGAGAGCTGGTGGCCCAAGAGATCAATTTCTCTGGATCAGAGTGAG 600  
DB 541 CATCGTAAGAGGAGAGCTGGTGGCCCAAGAGATCAATTTCTCTGGATCAGAGTGAG 600  
QY 601 GGGTCCCTGCTCGATGATCTTCTTGGATAACCCCTGATGATCTGAGATTAACCGTGAT 660

DB 601 GGGTCCCTGCTCGATGATCTTCTTGGATACCCCTGATGATGATTAACCGTGAT 660  
QY 661 GACATCGAGAGACCCCGATGAGACCGACTCGCTGGAGTTCTTGGGAAATGGCAACGAATG 720  
DB 661 GACATCGAGAGACCCCGATGAGACCGACTCGCTGGAGTTCTTGGGAAATGGCAACGAATG 720  
QY 721 GAGTGGGAGAGACGACACCCCGTGGCCACCGCAAGAACATGCCCCGGGAGACAGCGCGAT 780  
DB 721 GAGTGGGAGAGACGACACCCCGTGGCCACCGCAAGAACATGCCCCGGGAGACAGCGCGAT 780  
QY 781 CTATTTGGGAGCGGACGACGAGAGGACCGGAGCGCGCAACGGCGCTGTGGCGGACA 840  
DB 781 CTATTTGGGAGCGGACGACGAGAGGACCGGAGCGCGCAACGGCGCTGTGGCGGACA 840  
QY 841 GTGATCATCGGGGAGCAGACCGTATAGACCTGACATGATGATCCGCGCTTACATGAAA 900  
DB 841 GTGATCATCGGGGAGCAGACCGTATAGACCTGACATGATGATCCGCGCTTACATGAAA 900  
QY 901 GTGGTCAACCCACGGAGGCTTACCGGCGAAGGCTTCAACGGCATCATCGTCTTCGAGCC 960  
DB 901 GTGGTCAACCCACGGAGGCTTACCGGCGAAGGCTTCAACGGCATCATCGTCTTCGAGCC 960  
QY 961 TGTCTTCCAGAGCAGCAGCGCTCCCGACTACCACTACATCATGAGAAACCTCTTCCTG 1020  
DB 961 TGTCTTCCAGAGCAGCAGCGCTCCCGACTACCACTACATCATGAGAAACCTCTTCCTG 1020  
QY 1021 TACGTCATCAGCAGCTTACGCTCTGCTGGCTGAGGACTACATGATCGTGTACCTGTAAC 1080  
DB 1021 TACGTCATCAGCAGCTTACGCTCTGCTGGCTGAGGACTACATGATCGTGTACCTGTAAC 1080  
QY 1081 GGTGCCAGCCCGCGGAGGATGCTCGAATCGGCTGGCTGGAAGAGTGTCTACCAATG 1140  
DB 1081 GGTGCCAGCCCGCGGAGGATGCTCGAATCGGCTGGCTGGAAGAGTGTCTACCAATG 1140  
QY 1141 ATCGACCGGAGGTTGCGGAAAAAAGCTGAACTCTTGTATCATCGTCCACCCCTCGTGTTC 1200  
DB 1141 ATCGACCGGAGGTTGCGGAAAAAAGCTGAACTCTTGTATCATCGTCCACCCCTCGTGTTC 1200  
QY 1201 ATTGGGAGCTGTGGCCATCTCTCGCCCTTTCATCAGGCTCAGTTCATCAACAGATC 1260  
DB 1201 ATTGGGAGCTGTGGCCATCTCTCGCCCTTTCATCAGGCTCAGTTCATCAACAGATC 1260  
QY 1261 CAGTACGTGCAAGCTTGGAGAGCCTGGAGCAACTCATCCCTATGGAACAGTTCAGATC 1320  
DB 1261 CAGTACGTGCAAGCTTGGAGAGCCTGGAGCAACTCATCCCTATGGAACAGTTCAGATC 1320  
QY 1321 CAGACTGCTGCTGCAATACAGAGAGAAAGACTGAAAGGCCAGAGGAGAGCGCGAGG 1380  
DB 1321 CAGACTGCTGCTGCAATACAGAGAGAAAGACTGAAAGGCCAGAGGAGAGCGCGAGG 1380  
QY 1381 CCCAGCGGAGTTGTGCTGCCAGGTCGAAAGAGAGCCAGAGTGGCACAGTGGAA 1440  
DB 1381 CCCAGCGGAGTTGTGCTGCCAGGTCGAAAGAGAGCCAGAGTGGCACAGTGGAA 1440  
QY 1441 AACAGGCTGTCTGCTCAGAGATCAGGAAACAGCATGTCTCAGGCGCAGCTGAGC 1500  
DB 1441 AACAGGCTGTCTGCTCAGAGATCAGGAAACAGCATGTCTCAGGCGCAGCTGAGC 1500  
QY 1501 ATAAAGAGGACATGGAAGAGATTCAGATGCGCAAGAAAACCTCTGTCAAGCGCCACTG 1560  
DB 1501 ATAAAGAGGACATGGAAGAGATTCAGATGCGCAAGAAAACCTCTGTCAAGCGCCACTG 1560  
QY 1561 GCCCAGATCTATCTGCTCATCTCAGTCCCAATCTTCCAGGGTGGCAGCCCTCC 1620  
DB 1561 GCCCAGATCTATCTGCTCATCTCAGTCCCAATCTTCCAGGGTGGCAGCCCTCC 1620  
QY 1621 GTTCACTCTGAAAACCCAGCATCTCTTTTTCAGTGTCTGAAAACATTTGATTTTTTTTT 1680  
DB 1621 GTTCACTCTGAAAACCCAGCATCTCTTTTTCAGTGTCTGAAAACATTTGATTTTTTTTT 1680  
QY 1681 TAAAGATGAGATTTTGTGGCTTCAGAAAAGGGGCCAGTCTGAGCCCTCAGCCCTCC 1740

```
Db 1681 TAACGATCAGTATTTTGTGGTTCAGAAAAGGGCCAGCTCTGAGCCCTCCACCCCTTCC 1740
QY 1741 ACACTCAGAACTCTCAGCCGAGGAAGCAAGAGCGGAGGGGTGGCCCGGCTGCGGTC 1800
Db 1741 ACACTCAGAACTCTCAGCCGAGGAAGCAAGAGCGGAGGGGTGGCCCGGCTGCGGTC 1800
QY 1801 GGTGGCTCTGCTCTGCTGCGAGCCCTGTGTGTGAGAGTGAATCAAGATTCAGAGCC 1860
Db 1801 GGTGGCTCTGCTCTGCTGCGAGCCCTGTGTGTGAGAGTGAATCAAGATTCAGAGCC 1860
QY 1861 CTTCTCTGCTCTGCTCAGCCGCTCCAGGTTGGAGCCACAGACCCACCGCCACCGGCT 1920
Db 1861 CTTCTCTGCTCTGCTCAGCCGCTCCAGGTTGGAGCCACAGACCCACCGCCACCGGCT 1920
QY 1921 GGGTCTGCGCTCTTCTGCTGCTTCTCCCTCAGAAATGCGGCTCAGACCTAGAAAGCTCA 1980
Db 1921 GGGTCTGCGCTCTTCTGCTGCTTCTCCCTCAGAAATGCGGCTCAGACCTAGAAAGCTCA 1980
QY 1981 ACCCCCTATGAGGGCCAGCTCTGCGGTAGCTCTGACCTCGACCTTATGTCAAAT 2040
Db 1981 ACCCCCTATGAGGGCCAGCTCTGCGGTAGCTCTGACCTCGACCTTATGTCAAAT 2040
QY 2041 TCACACCCATGTTTTTCAATTTGACCCGCCCCCTTCTGCTCATATGACACCCAGCTCC 2100
Db 2041 TCACACCCATGTTTTTCAATTTGACCCGCCCCCTTCTGCTCATATGACACCCAGCTCC 2100
QY 2101 TTTGAGAGGATCAGAGCCCATTTGCAACAAGAGCGCGCTGCCAACCATCTGTCTCG 2160
Db 2101 TTTGAGAGGATCAGAGCCCATTTGCAACAAGAGCGCGCTGCCAACCATCTGTCTCG 2160
QY 2161 ATTGCAAAATGACACCCAGTAATCTAGAACTTTCTCAAGCCCTTTAACTCAGATGCA 2220
Db 2161 ATTGCAAAATGACACCCAGTAATCTAGAACTTTCTCAAGCCCTTTAACTCAGATGCA 2220
QY 2221 AGCCACCGGGCAAAACCCCGTCAATCTCCCAAGAGATGAGATATGTGACCTCACT 2280
Db 2221 AGCCACCGGGCAAAACCCCGTCAATCTCCCAAGAGATGAGATATGTGACCTCACT 2280
QY 2281 GCTCCCGGCAAAACCCCGTCAATCTCCCAAGAGATGAGATATGTGACCTCACT 2340
Db 2281 GCTCCCGGCAAAACCCCGTCAATCTCCCAAGAGATGAGATATGTGACCTCACT 2340
QY 2341 AGGCTCAGAAAATGAGTCTGAAATAGGAGAGCTTAGCAAGAGAGAGATACAGGATCGG 2400
Db 2341 AGGCTCAGAAAATGAGTCTGAAATAGGAGAGCTTAGCAAGAGAGAGATACAGGATCGG 2400
QY 2401 GGGTTGAGTCTTCAAGAGTCAATTCGGGAAGATTAATCCAGTGGCTGGCCGAGCCAC 2460
Db 2401 GGGTTGAGTCTTCAAGAGTCAATTCGGGAAGATTAATCCAGTGGCTGGCCGAGCCAC 2460
QY 2461 CTGCAATCAAGCTTGGACAGCGGTTCTGTTCCGGAGGCAAAATTTCCCTAGGAAAA 2520
Db 2461 CTGCAATCAAGCTTGGACAGCGGTTCTGTTCCGGAGGCAAAATTTCCCTAGGAAAA 2520
QY 2521 GAAGACAGACTTTTCTAATGTGTCGTAATGCGGATCACTGTGAGATGGAATCTAGAG 2580
Db 2521 GAAGACAGACTTTTCTAATGTGTCGTAATGCGGATCACTGTGAGATGGAATCTAGAG 2580
QY 2581 CACTGAGTCTGCTGCTGGAAGTATTAAGAAAAGGCTGGCCAGGCAAGATGGCTCA 2640
Db 2581 CACTGAGTCTGCTGCTGGAAGTATTAAGAAAAGGCTGGCCAGGCAAGATGGCTCA 2640
QY 2641 GCGCTGTAATCCAGACTTTGGAGCGCGAGCGGATCACTCTACTAAATAACAAAATTAGCA 2700
Db 2641 GCGCTGTAATCCAGACTTTGGAGCGCGAGCGGATCACTCTACTAAATAACAAAATTAGCA 2700
QY 2701 GAGAACAGCTGGCCCAACATGTGTGAACCTCTACTCTACTAAATAACAAAATTAGCA 2760
Db 2701 GAGAACAGCTGGCCCAACATGTGTGAACCTCTACTCTACTAAATAACAAAATTAGCA 2760
QY 2761 GCGCTGAGGAGGCTGCTGTAATCCAGTACTTGGAGGCTGAGGATGAGATCACT 2820
Db 2761 GCGCTGAGGAGGCTGCTGTAATCCAGTACTTGGAGGCTGAGGATGAGATCACT 2820
```

---

```
QY 2821 TAAACCTGAGGCGAGAGGTTACAGTGAGCCAGATCGTGCCACTGCAATTCAGGCTGG 2880
Db 2821 TAAACCTGAGGCGAGAGGTTACAGTGAGCCAGATCGTGCCACTGCAATTCAGGCTGG 2880
QY 2881 CGACAGACAAAGACTCTCTCTCAAAAAA 2918
Db 2881 CGACAGACAAAGACTCTCTCTCAAAAAA 2918

RESULT 2
ADRI6281
ID ADRI6281 standard; cDNA; 2918 BP.
XX AC ADRI6281;
XX DT 21-OCT-2004 (first entry)
XX DE Human Cayman ataxia variant cDNA.
XX KW Human; jittery; ss; gene; Cayman ataxia; ATCAY; KIAA1872;
XX KW Chromosome 19p13.3; ataxia; myoclonus; dystonia; epilepsy; myasthenia;
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
XX CDS 373..1488
XX FT /*tag= a
XX FT /product= "Cayman ataxia protein"
XX FT replace(1275,C)
XX FT /*tag= b
XX FT /standard_name= "single nucleotide polymorphism"
XX PN US2004146900-A1.
XX PD 29-JUL-2004.
XX PF 03-NOV-2003; 2003US-00699941.
XX PR 01-NOV-2002; 2002US-0422971P.
XX PR 08-NOV-2002; 2002US-0424973P.
XX PA (UNWI ) UNIV MICHIGAN.
XX PI Burmeister M;
XX DR WPI; 2004-552665/53.
XX DR P-PSDB; ADRI6282.
XX PT Detecting variant Cayman ataxia polypeptide or nucleic acid sequence in
XX PT subject, by detecting presence or absence of variant Cayman ataxia
XX PT polypeptide or nucleic acid in biological sample.
XX PS Claim 5; SEQ ID NO 8; 98pp; English.
XX CC The invention relates to detecting variant Cayman ataxia polypeptide or
XX CC nucleic acid sequence (ATCAY) in a subject, involving providing a
XX CC biological sample from a subject, where the biological sample comprises a
XX CC Cayman ataxia polypeptide or nucleic acid, and detecting the presence or
XX CC absence of a variant Cayman ataxia polypeptide or nucleic acid in the
XX CC biological sample. Also included is a kit comprising a reagent for
XX CC detecting the presence or absence of a variant Cayman ataxia nucleic acid
XX CC or polypeptide in a biological sample. The human ATCAY gene (previously
XX CC isolated as cDNA KIAA1872, located on chromosome 19p13.3) was identified
XX CC and mapped using the fact that it is the homologue of the mouse jittery
XX CC gene, located on mouse chromosome 10. The method is useful for detecting
XX CC variant Cayman ataxia polypeptide or nucleic acid sequence in a subject.
XX CC The presence of the variant Cayman ataxia polypeptide or nucleic acid is
XX CC indicative of a disorder chosen from ataxia, myoclonus, dystonia,
XX CC epilepsy, and myasthenia in the subject. The biological sample is chosen
XX CC from blood sample, a tissue sample, urine sample, saliva sample, and an
XX CC amniotic fluid sample. The subject is chosen from embryo, foetus, newborn
```





QY 2041 TCACACCCATGGTTTTCATTGACCCGCCGCCCTTCTGCTCATATGACACCCAGCTCC 2100  
DB |||||  
QY 2041 TCACACCCATGGTTTTCATTGACCCGCCGCCCTTCTGCTCATATGACACCCAGCTCC 2100  
DB |||||  
QY 2101 TTTGAGAGGATCAGAGCCCATTCACAGAGAGCGCTGCCAACCATCTTGTCTCTCG 2160  
DB |||||  
QY 2101 TTTGAGAGGATCAGAGCCCATTCACAGAGAGCGCTGCCAACCATCTTGTCTCTCG 2160  
DB |||||  
QY 2161 ATTGCAAAATGACACCCCGTCAATCTAGAACATTTCTCAAGCCCTTTAACTCAGATGTC 2220  
DB |||||  
QY 2161 ATTGCAAAATGACACCCCGTCAATCTAGAACATTTCTCAAGCCCTTTAACTCAGATGTC 2220  
DB |||||  
QY 2221 AGCCACGGGCAAAACCCCGTCAATCTAGAACATTTCTCAAGCCCTTTAACTCAGATGTC 2280  
DB |||||  
QY 2221 AGCCACGGGCAAAACCCCGTCAATCTAGAACATTTCTCAAGCCCTTTAACTCAGATGTC 2280  
DB |||||  
QY 2281 GCTCCCGCCCAACCCAGCGTCAAGGCTGGGACACGCGCAACGCTGTTCCGGTTGGAACAGCAG 2340  
DB |||||  
QY 2281 GCTCCCGCCCAACCCAGCGTCAAGGCTGGGACACGCGCAACGCTGTTCCGGTTGGAACAGCAG 2340  
DB |||||  
QY 2341 AGGCTCAGAACTGGCTCTGAAATAGGACAGCTAGCAAGAGAGAGATACAGGTTATCGG 2400  
DB |||||  
QY 2341 AGGCTCAGAACTGGCTCTGAAATAGGACAGCTAGCAAGAGAGAGATACAGGTTATCGG 2400  
DB |||||  
QY 2401 GCGTTTGAAGTTTCAAGAGTCAATTCGGAAGATTAATCCAGTGCCTGCGCGCAGCCAC 2460  
DB |||||  
QY 2401 GCGTTTGAAGTTTCAAGAGTCAATTCGGAAGATTAATCCAGTGCCTGCGCGCAGCCAC 2460  
DB |||||  
QY 2461 CTGCATTCAAGCTTGGACAGCGGCTTCTGTTGGGAGGCAATTTCCCTAGGAAAA 2520  
DB |||||  
QY 2461 CTGCATTCAAGCTTGGACAGCGGCTTCTGTTGGGAGGCAATTTCCCTAGGAAAA 2520  
DB |||||  
QY 2521 GAAGCAGACTTTCTAAATGTGTCCAAATGCGATCACTGTGTCAGATGATCTAGAG 2580  
DB |||||  
QY 2521 GAAGCAGACTTTCTAAATGTGTCCAAATGCGATCACTGTGTCAGATGATCTAGAG 2580  
DB |||||  
QY 2581 CACTGAGCTCCCTGCTCTGGAAGTATTTAAGAAAGGCTGGGCGCAGGACCATGGCTCA 2640  
DB |||||  
QY 2581 CACTGAGCTCCCTGCTCTGGAAGTATTTAAGAAAGGCTGGGCGCAGGACCATGGCTCA 2640  
DB |||||  
QY 2641 CGCCTGTAATCCAGACTTTGGAGCGCCAGGACGCGGATCACTGAGGTGAGGATTT 2700  
DB |||||  
QY 2641 CGCCTGTAATCCAGACTTTGGAGCGCCAGGACGCGGATCACTGAGGTGAGGATTT 2700  
DB |||||  
QY 2701 GAGACAGCTGCGCAACATGTGAACTCTATCTCTAATAAATACAAAATTAGCCA 2760  
DB |||||  
QY 2701 GAGACAGCTGCGCAACATGTGAACTCTATCTCTAATAAATACAAAATTAGCCA 2760  
DB |||||  
QY 2761 GCGGTGGTGGCAGGTGCTGTATATCCAGTACTTGGGAGGCTGAGGCAATGAGATCACT 2820  
DB |||||  
QY 2761 GCGGTGGTGGCAGGTGCTGTATATCCAGTACTTGGGAGGCTGAGGCAATGAGATCACT 2820  
DB |||||  
QY 2821 TAACTGAGAGCGCAGAGTTACAGTGACCCAGATCGGCCACTTCAGCTCAGCTGG 2880  
DB |||||  
QY 2821 TAACTGAGAGCGCAGAGTTACAGTGACCCAGATCGTGCCACTTCATTCAGCTGG 2880  
DB |||||  
QY 2881 CCACAGAGCAGACTGCTGCTCAAAAAA 2918  
DB |||||  
QY 2881 CCACAGAGCAGACTGCTGCTCAAAAAA 2918  
DB |||||

## RESULT 3

ADQ85643

ID ADQ85643 standard; cDNA; 4978 BP.

XX AC ADQ85643;

XX XX

XX DT 07-OCT-2004 (first entry)

XX XX

DE Human tumour-associated antigenic target (TAT) cDNA sequence #2457.

XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;  
KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;

KW cancer; cell proliferative disorder; gene; ss.  
XX Homo sapiens.  
XX WO2004060270-A2.  
XX 22-JUL-2004.  
XX 15-OCT-2003; 2003WO-US029126.  
XX 18-OCT-2002; 2002US-0418988P.  
XX (GETH ) GENENTECH INC.  
XX (WUTD/) WU T D.  
XX (ZHOU/) ZHOU Y.  
XX WU TD, Zhou Y;  
XX WPI; 2004-534300/51.  
XX New nucleic acid molecule and encoded polypeptide, for diagnosing, preventing or treating cell proliferative disorders such as cancer.  
XX Claim 1; SEQ ID NO 2457; 5504pp; English.

The present invention describes an isolated tumour-associated antigenic target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of (a); (c) the complement of (a) or (b); (d) a sequence that has 80% sequence identity to (a)-(c); or (e) a sequence that hybridizes to (a)-(c). Also described: (1) an expression vector comprising the above nucleic acid; (2) a host cell comprising the above expression vector; (3) a process for producing a polypeptide; (4) an isolated polypeptide comprising: (a) an amino acid sequence encoded by any of the above nucleotide sequences; (b) an amino acid sequence encoded by the full-length coding region of the above nucleotide sequences; or (c) a sequence having at least 80% identical to (a) or (b); (5) a chimeric polypeptide comprising the above polypeptide fused to a heterologous polypeptide; (6) an isolated antibody that binds to the above polypeptide; (7) a process for producing the antibody; (8) an isolated oligopeptide that binds to the above polypeptide; (9) a tumour-associated antigenic target (TAT) binding organic molecule that binds to the above polypeptide; (10) a composition of matter comprising the above (chimeric) polypeptide, antibody, oligopeptide or TAT binding organic molecule, in combination with a carrier; (11) an article of manufacture comprising a container and the composition of matter contained within the container; (12) methods of inhibiting the growth of a cell that expresses the above protein, where the growth of the cell is at least in part dependent upon a growth potentiating effect of the above protein; (13) a method of therapeutically treating a mammal having a cancerous tumour comprising cells that express the above protein; (14) a method of determining the presence of a protein in a sample suspected of containing the protein described above; (15) methods of diagnosing the presence of a tumour in a mammal; (16) a method for treating or preventing a cell proliferative disorder associated with increased expression or activity of the above protein; and (17) a method of binding an antibody, oligopeptide or organic molecule to a cell that expresses the protein described above. The TAT sequences have cytostatic activities, and can be used in gene therapy. The composition and methods are useful for diagnosing, preventing or treating cancer. The composition is also used for preparing a medicament for the therapeutic treatment or diagnostic detection of a cell proliferative disorder or cancer. The present sequence represents a human TAT cDNA sequence from the present invention.

Sequence 4978 BP; 1241 A; 1416 C; 1366 G; 955 T; 0 U; 0 Other;

## Query Match

Best Local Similarity 99.9%; Score 2913.8; DB 12; Length 4978;

Matches 2915; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCGAGCCTCTGCCAGCCTGAGCTGGGAAGAAGCAGCTACTCTCGAGCGGGGCCCA 60

DB 5 GCCGAGCCTCTGCCAGCCTGAGCTGGGAAGAAGCAGCTACTCTCGAGCGGGGCCCA 64





Db 2225 AGCCACCGGGGAAACCCCGTCAATACCTCCCAACAGGAATGAGATATGTGGACCTCACT 2284  
Qy 2281 GCTCCCCCAACCCAGCGTCAAGGCTGGGACACGCGCAACGCTGTTCGGGGTTGGAAACAGCAG 2340  
Db 2285 GCTCCCCCAACCCAGCGTCAAGGCTGGGACACGCGCAACGCTGTTCGGGGTTGGAAACAGCAG 2344  
Qy 2341 AGGCTCAGAACTGGCTCTGAATATAGGACAGCCTAGCAAGAGAAAGATACAGGGTATCGG 2400  
Db 2345 AGGCTCAGAACTGGCTCTGAATATAGGACAGCCTAGCAAGAGAAAGATACAGGGTATCGG 2404  
Qy 2401 GCGTTGAGGTGTTTCAGAGCTCAATCGGGAAGATAAATCAGTGCCTGCGCCAGCCAC 2460  
Db 2405 GCGTTGAGGTGTTTCAGAGCTCAATCGGGAAGATAAATCAGTGCCTGCGCCAGCCAC 2464  
Qy 2461 CTGCATTCAAAGCTTGACACAGCGGGTCTTGTTCGGGAGCAAAATTCCTTAGGAAAAA 2520  
Db 2465 CTGCATTCAAAGCTTGACACAGCGGGTCTTGTTCGGGAGCAAAATTCCTTAGGAAAAA 2524  
Qy 2521 GAAGACAGCTTTTCTAATGTGTCCAAATGCGGATCACTGGTCAGATGAGACTCTAGAAG 2580  
Db 2525 GAAGACAGCTTTTCTAATGTGTCCAAATGCGGATCACTGGTCAGATGAGACTCTAGAAG 2584  
Qy 2581 CACTGAGCTCCCTGCTCTGGAAGTATTTAAGAAAGGCTGGGCCAGGACAGATGCTCA 2640  
Db 2585 CACTGAGCTCCCTGCTCTGGAAGTATTTAAGAAAGGCTGGGCCAGGACAGATGCTCA 2644  
Qy 2641 CGCTCTAATCCAGACTTTGGGAGCGCGAGGCGGATCACCTGAGTGAAGGATTT 2700  
Db 2645 CGCTCTAATCCAGACTTTGGGAGCGCGAGGCGGATCACCTGAGTGAAGGATTT 2704  
Qy 2701 GAGAACAGCTGCGCCAACTGTTGAACCTCTCTACTTAAATAATACAAAATAGCCA 2760  
Db 2705 GAGAACAGCTGCGCCAACTGTTGAACCTCTCTACTTAAATAATACAAAATAGCCA 2764  
Qy 2761 GCGTGTGGCAGGCTGCTGTATCCAGCTACTTGGAGGCTGAGGCGATGAGATCAT 2820  
Db 2765 GCGTGTGGCAGGCTGCTGTATCCAGCTACTTGGAGGCTGAGGCGATGAGATCAT 2824  
Qy 2821 TAAACCTGAGAGCGAGGTTACAGTGAGCGCAAGATCGTGCCACTGCATTCCAGCTGGG 2880  
Db 2825 TAAACCTGAGAGCGAGGTTACAGTGAGCGCAAGATCGTGCCACTGCATTCCAGCTGGG 2884  
Qy 2881 CGACAGACGAGACTCTGTCTCAAAAAAATAAAAAA 2917  
Db 2885 CGACAGACGAGACTCTGTCTCAAAAAAATAAAAAA 2921

RESULT 4

ACN40413  
ID ACN40413 standard; cDNA; 4978 BP.  
XX ACN40413;  
AC ACN40413;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Tumour-associated antigenic target (TAT) cDNA DNA326528, SEQ ID NO:5176.  
XX  
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;  
KW tumour; diagnosis; cell proliferative disorder; breast cancer;  
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;  
KW central nervous system cancer; bladder cancer; pancreatic cancer;  
KW cervical cancer; melanoma; leukaemia; hybridisation probe;  
KW chromosome identification; chromosome mapping; gene mapping;  
KW gene therapy; cytostatic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2004030615-A2.  
XX  
PD 15-APR-2004.  
XX  
PF 29-SEP-2003; 2003WO-US028547.

XX  
PR 02-OCT-2002; 2002US-0414971P.  
XX (GETH ) GENENTECH INC.  
XX Wu TD, Zhang Z, Zhou Y;  
PI WPI; 2004-347921/32.  
DR P-PSDB; ABM82007.  
XX  
PT New tumor-associated antigenic target polypeptides and nucleic acids,  
PT useful in preparing a medicament for treating or detecting a  
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or  
PT prostate cancer or tumor.  
XX  
PS Claim 1; SEQ ID NO 5176; 7273pp; English.  
XX  
CC The invention relates to human tumour-associated antigenic target (TAT)  
CC polypeptides, and their related nucleic acids. The TAT polypeptides are  
CC overexpressed in cancer tissues compared to normal tissues, and may thus  
CC serve as effective targets for the diagnosis and treatment of cancer in  
CC mammals. The invention also relates to nucleic acid and polypeptide  
CC sequences at least 80% identical to the TAT nucleic acids and  
CC polypeptides; expression vectors and host cells comprising a TAT nucleic  
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic  
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a  
CC TAT polypeptide; and methods and compositions for the treatment or  
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,  
CC antibodies, antagonists, binding molecules and compositions are useful  
CC for diagnosing or treating a cell proliferative disorder associated with  
CC increased TAT expression, particularly cancers such as breast cancer,  
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder  
CC cancer, pancreatic cancer, cervical cancer, cancers of the central  
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be  
CC used as hybridisation probes, in chromosome and gene mapping, in  
CC chromosome identification and in gene therapy. The present sequence  
CC represents a TAT nucleic acid of the invention  
XX  
SQ Sequence 4978 BP; 1241 A; 1416 C; 1366 G; 955 T; 0 U; 0 Other;

Query Match 99.9%; Score 2913.8; DB 13; Length 4978;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2915; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 GCCGAGCTCTGCGAGCCCTGAGCTGGGAAGAGAGCTACTCGGAGCAGGGCGCGCA 60  
Db 5 GCCGAGCTCTGCGAGCCCTGAGCTGGGAAGAGAGCTACTCGGAGCAGGGCGCGCA 64  
Qy 61 GCGCGGCGCGATGAGAGGGGGCGCAGCGCGAGCCCGGCTGGGGGCGCCACCGCTAAC 120  
Db 65 GCGCGGCGCGATGAGAGGGGGCGCAGCGCGAGCCCGGCTGGGGGCGCCACCGCTAAC 124  
Qy 121 CTTGACACCCACCCACCCCTTCACAAAAGAGCTGGCGGGCGCTGGCCACGTCGCGCTGGG 180  
Db 125 CTTGACACCCACCCACCCCTTCACAAAAGAGCTGGCGGGCGCTGGCCACGTCGCGCTGGG 184  
Qy 181 TGACCTTCTCGATGACAGATTCGCCCTTCGAGAGATCTCTTCTCTTAGGCTCTGAA 240  
Db 185 TGACCTTCTCGATGACAGATTCGCCCTTCGAGAGATCTCTTCTCTTAGGCTCTGAA 244  
Qy 241 GCGCGGAGGCGTGAGCGATGCCAGCTGCACCGGCGAGGCTCGCTTTGTTGCA 300  
Db 245 GCGCGGAGGCGTGAGCGATGCCAGCTGCACCGGCGAGGCTCGCTTTGTTGCA 304  
Qy 301 GTAAGGAGGAGGCTGTCTCAGCTGACAGGGGTATCCCTGCTTCAAGCAGGCTCTC 360  
Db 305 GTAAGGAGGAGGCTGTCTCAGCTGACAGGGGTATCCCTGCTTCAAGCAGGCTCTC 364  
Qy 361 TTCCAGCTCCCATGGGACACCGAAGCCACGCTCCGGATGGAAAACTGAGAGCTGAAG 420  
Db 365 TTCCAGCTCCCATGGGACACCGAAGCCACGCTCCGGATGGAAAACTGAGAGCTGAAG 424  
Qy 421 GAGGATGGCAGGAGGAGATCTTCCAGGCCATCTCCCAAGAGAGACGCGGGGTGGAAC 480

	Db	 GAGGAATGGCAGGACGAAGATCTTCCCAAGGCCACTCCACAGAAGAGACGGGGGTGGAACTG	484
	Qy	CTTGCGACGCCGGTGGGAAGACAATCCTCTCTCCTCCCAACGCCTAATAATTTTCACAGCGAGCG	540
	Db	CTTGCGACGCCGGTGGGAAGACAATCCTCTCTCTCTCCCAACGCCTAATAATTTTCACAGCGAGCG	544
	Qy	CATCGTAAGAGAGAAAGACGCTGGTGGCCCCCAGAGATCAACATTTCTCTGGATCAGAGTGAG	600
	Db	CATCGTAAGAGAGAAAGACGCTGGTGGCCCCCAGAGATCAACATTTCTCTGGATCAGAGTGAG	604
	Qy	GGGTCCCTGCTGTCGGATGACATTCTTGGATACCCTGATGACCTGGATATTAACGCTGGAT	660
	Db	GGGTCCCTGCTGTCGGATGACATTCTTGGATACCCTGATGACCTGGATATTAACGCTGGAT	664
	Qy	GACATCGAGACCCCCTGATGACACGACTCGCTGGAGTTCTCTGGGGAATGGCAACGAAC TG	720
	Db	GACATCGAGACCCCCTGATGACACGACTCGCTGGAGTTCTCTGGGGAATGGCAACGAAC TG	724
	Qy	GAGTGGGAAGACGACACCCCCGTGGCCACCGCCCAAGAACATSCCCGGGGACAGCGCGGAT	780
	Db	GAGTGGGAAGACGACACCCCCGTGGCCACCGCCCAAGAACATSCCCGGGGACAGCGCGGAT	784
	Qy	CTATTGTTGGGACCGGACA CGACGGAGGA CCGCAGCGCGCCAAACGGCGCTGTGTGCGGACA	840
	Db	CTATTGTTGGGACCGGACA CGACGGAGGA CCGCAGCGCGCCAAACGGCGCTGTGTGCGGACA	844
	Qy	GTGATCATCGGGGAGCAAGACCGGTATAGACCTGCACATGATCCGGCCCTTACATGAAA	900
	Db	GTGATCATCGGGGAGCAAGACCGGTATAGACCTGCACATGATCCGGCCCTTACATGAAA	904
	Qy	GTGGTCAACCCACGGAGGGTACTACGGCGGAGGECCTCAAACGCATCATCTCTCGCAGCC	960
	Db	GTGGTCAACCCACGGAGGGTACTACGGCGGAGGECCTCAAACGCATCATCTCTCGCAGCC	964
	Qy	TGCTTCTCTTCAGACAGCAGGCTCCCGGAC TPA CCA CTACATCATGTGAGAAACCTCTTCTCTG	1020
	Db	TGCTTCTCTTCAGACAGCAGGCTCCCGGAC TPA CCA CTACATCATGTGAGAAACCTCTTCTCTG	1024
	Qy	TACGTCAATCACACGCTTAGAGCTCCTGGTGGCTCAGAGCTACATGATCGGTACTCTGAAC	1080
	Db	TACGTCAATCACACGCTTAGAGCTCCTGGTGGCTCAGAGCTACATGATCGGTACTCTGAAC	1084
	Qy	GGTGCCACGCCCCGGCGAGGATGCCTGGAATCGGCTGGCTGGAAGAAAGTGCTACACAGATG	1140
	Db	GGTGCCACGCCCCGGCGAGGATGCCTGGAATCGGCTGGCTGGAAGAAAGTGCTACACAGATG	1144
	Qy	ATCGACGGAGGTTGGGAAAAA CCTGAAAGTCTTGATCATCTGTCACCCCTCGTGGTTC	1200
	Db	ATCGACGGAGGTTGGGAAAAA CCTGAAAGTCTTGATCATCTGTCACCCCTCGTGGTTC	1204
	Qy	ATTCGGAGTGTGCTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTTCATCAACAGATC	1260
	Db	ATTCGGAGTGTGCTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTTCATCAACAGATC	1264
	Qy	CAGTACGTGCACAGCTTGGAAAGACCTGGAGCAACTCATCCCTATGGAACACGTCACAGATC	1320
	Db	CAGTACGTGCACAGCTTGGAAAGACCTGGAGCAACTCATCCCTATGGAACACGTCACAGATC	1324
	Qy	CCAGACTGTGCTCTGCAATACGAAGAGGAAAGACTGAAGGC CAGGAGGAGAGCGCAGG	1380
	Db	CCAGACTGTGCTCTGCAATACGAAGAGGAAAGACTGAAGGC CAGGAGGAGAGCGCAGG	1384
	Qy	CCCCCAGCGGAGTTGTGCTGCCAGGCTCTGAAGAGAGCCAGNAGTGGCCACCGTGGAA	1440
	Db	CCCCCAGCGGAGTTGTGCTGCCAGGCTCTGAAGAGAGCCAGNAGTGGCCACCGTGGAA	1444
	Qy	AACAGGCTGTGCTCTGGTCTCAGAAAGATCAGGAAAAAAGCATGTCTCAGGCGCAGCTGAGC	1500
	Db	AACAGGCTGTGCTCTGGTCTCAGAAAGATCAGGAAAAAAGCATGTCTCAGGCGCAGCTGAGC	1504
	Qy	ATAACAAAGGACATGGGAAGAAATTCAGATGCCAGAAAAACCTCTGTGACAGCCCCACTG	1560





PH	Key	Location/Qualifiers	
FT	5'UTR	1..369	
FT	CDS	/*tag= a	
FT		370..1485	
FT		/*tag= b	
FT		/product= "Human NIP2b protein"	
FT		/note= "this region is specifically claimed as	
FT		Seq.ID.NO:3"	
FT		1486..3065	
FT	3'UTR	/*tag= c	
XX			
PN	WO200023470-A1.		
XX			
XX	27-APR-2000.		
PD			
XX	19-OCT-1999;	99WO-US024335.	
PF			
XX	19-OCT-1998;	98US-00174937.	
PR			
XX	(MILL-) MILLENNIUM PHARM INC.		
PA			
PI	Curtis RAJ, Glucksmann MA;		
XX			
PI	WPI: 2000-339653/29.		
DR	P-PSDB; AAY70847.		
XX			
XX	Nucleic acids encoding NIP2 proteins family members designated NIP2b, NIP2cL and NIP2cS, useful for preventing, diagnosing and treating disorders associated with inappropriate NIP2 expression.		
PT			
PT			
XX			
PS	Claim 1a; Fig 1; 120pp; English.		
XX			
CC	The present sequence is the cDNA encoding the human NIP2b protein, a member of the NIP2 protein family. It is predominantly expressed in the brain and also in prostate and small intestine. It plays a major role in apoptosis. The NIP2b nucleotide sequence is 74% identical to human mRNA for KIAA0367. NIP2b interacts with proteins like adenoviral B1B 19 kDa protein or BCL-2 protein to provide a cell survival function. It activates NIP2b-dependent signal transduction pathway and modulates programmed cell death. This sequence can be used to screen for NIP2b, NIP2cL and NIP2cS sequences and mutations in them, in diagnostic and prognostic assays, pharmacogenetics and for therapeutic and prophylactic purposes. It can also be used to screen for naturally occurring substrates, drugs or modulators of their activity. The proteins can be used to treat disorders associated with programmed cell death		
CC			
XX	Sequence 3065 BP; 738 A; 906 C; 839 G; 582 T; 0 U; 0 Other;		
SQ			
	Query Match	99.0%; Score 2887.6; DB 3; Length 3065;	
	Best Local Similarity	99.9%; Pred. No. 0;	
	Matches 2890; Conservative	0; Mismatches 4; Indels 0; Gaps 0;	
QY	24	CTGGGAAGAGCAGCTACCTCGAGCAGGGGGCGGAGGGGGCGGCGGCGGAGAGGGGGC	83
DB	21	CGGGGAAGAGCAGCTACCTCGAGCAGGGGGCGGAGGGGGCGGCGGCGGAGAGGGGGC	80
QY	84	GCAGCGCAGCGCCCGCGCTGGGGAGCCACCGCTAACCTGACCCCGACCCACCCCTGCA	143
DB	81	GCAGCGCAGCGCCCGCGCTGGGGAGCCACCGCTAACCTGACCCCGACCCCGCTGCA	140
QY	144	CAAAAGAGTGCGGGCGCTGCGCCACGTCGCCCTGGGTGACCTTCTCGGATGACAGATC	203
DB	141	CAAAAGAGTGCGGGCGCTGCGCCACGTCGCCCTGGGTGACCTTCTCGGATGACAGATC	200
QY	204	CGCCCTGCGAGATCTCTTCTCTAGGCTCTGAGAGCCCGGGAGCGGTGAGCGATGC	263
DB	201	CGCCCTGCGAGATCTCTTCTCTAGGCTCTGAGAGCCCGGGAGCGGTGAGCGATGC	260
QY	264	CCAGCTGCACCCGGGCGAGGGCTCGCTTGTGTCAGTAAGGAGGAGGGCTGTCAG	323
DB	261	CCAGCTGCACCCGGGCGAGGGCTCGCTTGTGTCAGTAAGGAGGAGGGCTGTCAG	320
QY	324	CTGCAGAGGGGTCACTCCCTGCTTCAAGCCAGTGCCTCTTCCAGCTCCCATGGGAGCCAC	383

DB	321	CTGCAGAGGGGTCACTCCCTGCTTCAAGCCAGTGCCTTCCAGCTCCCATGGGAGCCAC	380
QY	384	CGAAGCCAGCTCCGGATGGAAAACGTGGAGCTGAAGAGGAGATGGCAGGACGAGATCT	443
DB	381	CGAAGCCAGCTCCGGATGGAAAACGTGGAGCTGAAGAGGAGATGGCAGGACGAGATCT	440
QY	444	TCCAGAGGCACCTCCCAAGAGAGACGGGGGTGGAACCTGCTTGGCAGCCCGGTGGAAAGAC	503
DB	441	TCCAGAGGCACCTCCCAAGAGAGACGGGGGTGGAACCTGCTTGGCAGCCCGGTGGAAAGAC	500
QY	504	ATCCTCTCTCTCCAAACAGCTTAAATTTCAAGAGCGCATCGTAAGAGAAACAGCTGCT	563
DB	501	ATCCTCTCTCTCCAAACAGCTTAAATTTCAAGAGCGCATCGTAAGAGAAACAGCTGCT	560
QY	564	GGCCCCAGAGATCAACATTTCTCTCGATCAGAGTGGGGTCCCTGCTGTCGATGACTT	623
DB	561	GGCCCCAGAGATCAACATTTCTCTCGATCAGAGTGGGGTCCCTGCTGTCGATGACTT	620
QY	624	CTTGGATACCCCTGATGACCTGGATATTAACGTGATGACATCGAGACCCCGATGAGAC	683
DB	621	CTTGGATACCCCTGATGACCTGGATATTAACGTGATGACATCGAGACCCCGATGAGAC	680
QY	684	CGACTCGCTGGAGTTCCTGGGGAATGGCAACGAACTGGAGTGGGAAGACGACACCCCGT	743
DB	681	CGACTCGCTGGAGTTCCTGGGGAATGGCAACGAACTGGAGTGGGAAGACGACACCCCGT	740
QY	744	GGCCACCCGCAAGACATCGCCGGGACAGCGCGGATCTATTGGGGAGCGGACACGAG	803
DB	741	GGCCACCCGCAAGACATCGCCGGGACAGCGCGGATCTATTGGGGAGCGGACACGAG	800
QY	804	GGAGCGCAGCGCCGCAACGGCGCTCTGGCGGACAGTGATCACTGGGGAGCAAGAGCA	863
DB	801	GGAGCGCAGCGCCGCAACGGCGCTCTGGCGGACAGTGATCACTGGGGAGCAAGAGCA	860
QY	864	CGGTATAGACCTGCAATGATCGGCTTACATGAAAGTGGTCACTCCAGGAGGGTACTA	923
DB	861	CGGTATAGACCTGCAATGATCGGCTTACATGAAAGTGGTCACTCCAGGAGGGTACTA	920
QY	924	CGGGAAGCGCTCAAGCCATCATGCTTTCGAGCGCTGCTTCTTCCAGACAGAGCT	983
DB	921	CGGGAAGCGCTCAAGCCATCATGCTTTCGAGCGCTGCTTCTTCCAGACAGAGCT	980
QY	984	CCCCGACTACCACTACATCATGAGAGAACTCTTCTGCTACGTCATCAGCAGCTTAGAGCT	1043
DB	981	CCCCGACTACCACTACATCATGAGAGAACTCTTCTGCTACGTCATCAGCAGCTTAGAGCT	1040
QY	1044	CCTGTGCTGAGGACTACATGATCGTGTACCTGAAACGGTGCACGCCCGCGCGAGGAT	1103
DB	1041	CCTGTGCTGAGGACTACATGATCGTGTACCTGAAACGGTGCACGCCCGCGCGAGGAT	1100
QY	1104	GCCTGGAATCGGCTGGCTGAAAGAGTGTCTACAGATGATGACCGAGGTTGCGGAAAA	1163
DB	1101	GCCTGGAATCGGCTGGCTGAAAGAGTGTCTACAGATGATGACCGAGGTTGCGGAAAA	1160
QY	1164	CCTGAAGTCTTGTATCATCGCTCCACCCCTCGTGGTTCATTGCGGATGCTGGCCATCTC	1223
DB	1161	CCTGAAGTCTTGTATCATCGCTCCACCCCTCGTGGTTCATTGCGGATGCTGGCCATCTC	1220
QY	1224	TGCGCTTTTCATCAGCGGTCAAGTTTCATCAACAGATTCAGTACGTGCAACAGCTTGAAGA	1283
DB	1221	TGCGCTTTTCATCAGCGGTCAAGTTTCATCAACAGATTCAGTACGTGCAACAGCTTGAAGA	1280
QY	1284	CCTGAGCACTCATATCCCTATGGAACAGCTCCAGATCCAGATGCTGCTGCTGCAATACGA	1343
DB	1281	CCTGAGCACTCATATCCCTATGGAACAGCTCCAGATCCAGATGCTGCTGCTGCAATACGA	1340
QY	1344	AGAGGAGACTGAGGCGCAGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG	1403
DB	1341	AGAGGAGACTGAGGCGCAGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG	1400
QY	1404	CAGGTCTGAAGAGAGCGCAGAGGTGGCACAGTGGAAACAGGTCTGCTGCTGCTCAG	1463



Db 1401 CAGGCTGTAAGAGAGCAGCAGAGGTGGCA CAGGTGGA AAAACAGGCTGCTGTGCTCTCAGA 1460  
Qy 1464 AGATCAGGAAACAGCATGTCTGAGCGACGTCAGCATAA CAAAGGACATGGAAGAGA 1523  
Db 1461 AGATCAGGAAACAGCATGTCTGAGCGACGTCAGCATAA CAAAGGACATGGAAGAGA 1520  
Qy 1524 TTCAGATGCCAGAAAACCTCTGTCTCAGAGCCCACTGAGCCCAAGATCTCATCTGCTCTCA 1583  
Db 1521 TTCAGATGCCAGAAAACCTCTGTCTCAGAGCCCACTGAGCCCAAGATCTCATCTGCTCTCA 1580  
Qy 1584 TCCTGAGTCCCAATCTTCCAAAGGGTCCAGCCCTCCGTTCTCTCTGAAACCCAGCATC 1643  
Db 1581 TCCTGAGTCCCAATCTTCCAAAGGGTCCAGCCCTCCGTTCTCTCTGAAACCCAGCATC 1640  
Qy 1644 CTTTTCAGCTGCTTGA AAAACATTTGATTTTTTTTTTTTAAACGATGCAATTTTGTGCGTT 1703  
Db 1641 CTTTTCAGCTGCTTGA AAAACATTTGATTTTTTTTTTTTAAACGATGCAATTTTGTGCGTT 1700  
Qy 1704 CAGAAAAGGGCCAGCTCTGAGCCCTCAACCTTCCACATCAAGAACTCTCAGCCGAG 1763  
Db 1701 CAGAAAAGGGCCAGCTCTGAGCCCTCAACCTTCCACATCAAGAACTCTCAGCCGAG 1760  
Qy 1764 GAAGGCAAGAGCGCAGGGGGTGGCCGGTGGCGTGGGCTCGGCTCCGCTCTGCTCGCA 1823  
Db 1761 GAAGGCAAGAGCGCAGGGGGTGGCCGGTGGCGTGGGCTCCGCTCTGCTCGCA 1820  
Qy 1824 GCCCTCTGTGTCTCAGAGCTGATCAAGATTTCAAGACCTTCTCTGCTTTGTCAACCGCTC 1883  
Db 1821 GCCCTCTGTGTCTCAGAGCTGATCAAGATTTCAAGACCTTCTCTGCTTTGTCAACCGCTC 1880  
Qy 1884 CAGGTTGAGGCCACAGACCCACCCCGCTGGGTCTGGCTCTCTTCTGTGCC 1943  
Db 1881 CAGGTTGAGGCCACAGACCCACCCCGCTGGGTCTGGCTCTCTTCTGTGCC 1940  
Qy 1944 TTTCCCTCCAGATCGGGCTCAGACCTAGAGCTCAACCCCGCTTATGAGGGCCAGCTCC 2003  
Db 1941 TTTCCCTCCAGATCGGGCTCAGACCTAGAGCTCAACCCCGCTTATGAGGGCCAGCTCC 2000  
Qy 2004 TGGGGTAGCTCTCGACCTCCGACCTTATGTCCAAATTTTCAACCCCAATGTTTTTANTTG 2063  
Db 2001 TGGGGTAGCTCTCGACCTCCGACCTTATGTCCAAATTTTCAACCCCAATGTTTTTANTTG 2060  
Qy 2064 ACCCGCCCTTCTCGCTCATATGACACCCAGCTCTTTGAGAGGATCAGAGCCCATTTG 2123  
Db 2061 ACCCGCCCTTCTCGCTCATATGACACCCAGCTCTTTGAGAGGATCAGAGCCCATTTG 2120  
Qy 2124 CACAGAGAGCGCTGCCACCATCTTGTCTCCGATGCAAAATGACACCCCGATTA 2183  
Db 2121 CACAGAGAGCGCTGCCACCATCTTGTCTCCGATGCAAAATGACACCCCGATTA 2180  
Qy 2184 TCTAGAACATTTCTCAAGCCCTTTTAACTCAGATGTCAAGCCACCGGGCAAAACCCGTCAA 2243  
Db 2181 TCTAGAACATTTCTCAAGCCCTTTTAACTCAGATGTCAAGCCACCGGGCAAAACCCGTCAA 2240  
Qy 2244 TACCTCCCAACCAAGGAATGAGATATGTGACCTCACTGTCTCCCCCAACCCAGCGTCAGGC 2303  
Db 2241 TACCTCCCAACCAAGGAATGAGATATGTGACCTCACTGTCTCCCCCAACCCAGCGTCAGGC 2300  
Qy 2304 TGGGACACCCACCGCTGTTCCGGTTGGACAGCAGAGGCTCAGAACTGCTCTGAAA 2363  
Db 2301 TGGGACACCCACCGCTGTTCCGGTTGGACAGCAGCAGGCTCAGAACTGCTCTCTGAAA 2360  
Qy 2364 TAGGCAGACCTTAGCAAGAGAGATACAGGTTATCGGGGTTTGTAGTGTTCAGAGTCA 2423  
Db 2361 TAGGCAGACCTTAGCAAGAGAGATACAGGTTATCGGGGTTTGTAGTGTTCAGAGTCA 2420  
Qy 2424 TTCGGGAAGATAAATCCAGTGGCTGGCCGACGACCTTGCATTTCAAAGCTTTGGACCCAGC 2483  
Db 2421 TTCGGGAAGATAAATCCAGTGGCTGGCCGACGACCTTGCATTTCAAAGCTTTGGACCCAGC 2480  
Qy 2484 GGGTTCTCTGTCGGGAGGCAAAATTTCCCTAGGAAAAGAGACAGACTTTTCTAATGTGG 2543  
Db 2481 GGGTTCTCTGTCGGGAGGCAAAATTTCCCTAGGAAAAGAGACAGACTTTTCTAATGTGGG 2540

Qy 2544 TCCAAATCGGATCACTGCTCAGATGGACTCTAGAAGCACTCAGCTCCCTGTCTCTGGAA 2603  
Db 2541 TCCAAATCGGATCACTGCTCAGATGGACTCTAGAAGCACTCAGCTCCCTGTCTCTGGAA 2600  
Qy 2604 GTATTTAAGAAAAGGCTGGCCAGGCACGATGGCTCACGCTGTATATCCAGACTTTGGG 2663  
Db 2601 GTATTTAAGAAAAGGCTGGCCAGGCACGATGGCTCACGCTGTATATCCAGACTTTGGG 2660  
Qy 2664 AGGCCGAGCGAGCGGCGATCACTGAGGTGAGGAGTTTGAGAACAGCCTGGCCCAATGGT 2723  
Db 2661 AGGCCGAGCGAGCGGCGATCACTGAGGTGAGGAGTTTGAGAACAGCCTGGCCCAATGGT 2720  
Qy 2724 GAAACCTCATCTCTACTATAAAAATCAAAAATTAGCCAGCGTGTGGCAGGTGCCGTAA 2783  
Db 2721 GAAACCTCATCTCTACTATAAAAATCAAAAATTAGCCAGCGTGTGGCAGGTGCCGTAA 2780  
Qy 2784 TCCAGACTACTTGGGAGGCTGAGGCATGAGAAATCACTTTAAACCTGAGAGGCGAGGTTAC 2843  
Db 2781 TCCAGACTACTTGGGAGGCTGAGGCATGAGAAATCACTTTAAACCTGAGAGGCGAGGTTAC 2840  
Qy 2844 AGTGAGCCAAAGTCTGTCACCTGCAATTCAGGCTGGGCGACAGACCAAGACTCTGTCTCA 2903  
Db 2841 AGTGAGCCAAAGTCTGTCACCTGCAATTCAGGCTGGGCGACAGACCAAGACTCTGTCTCA 2900  
Qy 2904 AAAAAA AAAAAA 2917  
Db 2901 AAAAAA AAAAAA 2914

## RESULT 7

AAH16751

ID AAH16751 standard; cDNA; 2894 BP.

XX AAH16751;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:15950.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 95JP-00248036.

XX 27-AUG-1999; 95JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX Claim 8; SEQ ID NO 15950; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the





1808 TCCGCTCTCTGCTGCGAGCCCTCTGTGCTCAGAGCTGGATACAAGATTCAAGACCCCTTCTCT 1867  
1800 TCCGCTCTCTGCTGCGAGCCCTCTGTGCTCAGAGCTGGATACAAGATTCAAGACCCCTTCTCT 1859  
1868 TGCCTGCTCAGCCGCTCCAGGTTGGAGCCACAGACACCCACCCGCTGCGGTCTG 1927  
1860 TGCCTGCTCAGCCGCTCCAGGTTGGAGCCACAGACACCCACCCGCTGCGGTCTG 1919  
1928 CGTCTCTTCTGCTGCTCTTCCCTCCAGAAATGCGGCTCAGACCTAGAAGCTCAACCCCTC 1987  
1920 CGTCTCTTCTGCTGCTCTTCCCTCCAGAAATGCGGCTCAGACCTAGAAGCTCAACCCCTC 1979  
1988 TATGAGGCGCAGCTCTGCGGTAGCTCTGACCTCCGACCTTATGTGCCAAATTTCAACCC 2047  
1980 TATGAGGCGCAGCTCTGCGGTAGCTCTGACCTCCGACCTTATGTGCCAAATTTCAACCC 2039  
2048 CATGTTTTTCTATTTACCGCCCTCTCTGCTCATATGACACCCAGCTCTTTGAGA 2107  
2040 CATGTTTTTCTATTTGACCGCCCTCTCTGCTCATATGACACCCAGCTCTTTGAGA 2099  
2108 GGATCAGAGCCCATTTGCAAGAAGAGCGCTGCCAAACCATCTCTGCTCTCGATTGCAA 2167  
2100 GGATCAGAGCCCATTTGCAAGAAGAGCGCTGCCAAACCATCTCTGCTCTCGATTGCAA 2159  
2168 ATGACACCCCGAGTAATCTAGAAATCTTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2227  
2160 ATGACACCCCGAGTAATCTAGAAATCTTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2219  
2228 GGGCAACCCCGTCAATACCTCCCAAGAGGATGAGATATGTGGACCTCACTGTCTCCC 2287  
2220 GGGCAACCCCGTCAATACCTCCCAAGAGGATGAGATATGTGGACCTCACTGTCTCCC 2279  
2288 CAAACCCAGGCTCAGGCTGGGACACGCGCTTCCGGTGGAAACAGAGAGCTCA 2347  
2280 CAAACCCAGGCTCAGGCTGGGACACGCGCTTCCGGTGGAAACAGAGAGCTCA 2339  
2348 GAACTGGCTCTGAATATGAGCAGCTTAGCAGAGGAAGATACAGGTTATCGGGGTTG 2407  
2340 GAACTGGCTCTGAATATGAGCAGCTTAGCAGAGGAAGATACAGGTTATCGGGGTTG 2399  
2408 AGTGTCTCAGAGTCAATTCGGGAGATTAATCCAGTGGCTGGCCGACGACCTGCAAT 2467  
2400 AGTGTCTCAGAGTCAATTCGGGAGATTAATCCAGTGGCTGGCCGACGACCTGCAAT 2459  
2468 CAAAGCTTGGACCCAGCGGTTCTTGTTCGGAGGCAATTTCCCTAGGAAAGAGACA 2527  
2460 CAAAGCTTGGACCCAGCGGTTCTTGTTCGGAGGCAATTTCCCTAGGAAAGAGACA 2519  
2528 GACTTTTCTAAATGTGTCCAAATGCGGATCACTGGTCAAGATGCACTGAGCACTGAG 2587  
2520 GACTTTTCTAAATGTGTCCAAATGCGGATCACTGGTCAAGATGCACTGAGCACTGAG 2579  
2588 CTCCTGCTCTGGAAGTATTTAAGAAAGCTTGGCCAGGCGACGATGCTCAGCCTGT 2647  
2580 CTCCTGCTCTGGAAGTATTTAAGAAAGCTTGGCCAGGCGACGATGCTCAGCCTGT 2639  
2648 AATCCAGACTTTGGAGGCGGAGGACGCGATCACTGAGTGGAGGATTTGAGAAC 2707  
2640 AATCCAGACTTTGGAGGCGGAGGACGCGATCACTGAGTGGAGGATTTGAGAAC 2699  
2708 GCTTGGCCAACTGGTGAACCTCATCTCTACTTAAATAATACAAATAATGAGCCAGCGTGG 2767  
2700 GCTTGGCCAACTGGTGAACCTCATCTCTACTTAAATAATACAAATAATGAGCCAGCGTGG 2759  
2768 TGGCAGTCTGCTTAATCCAGTCACTTGGAGGCTGAGGATGAGCAATCACTTAACT 2827  
2760 TGGCAGTCTGCTTAATCCAGTCACTTGGAGGCTGAGGATGAGCAATCACTTAACT 2819  
2828 GAGAGCAGAGGTTACAGTGAGCAAGATCGTGCCACTGCTTCCAGCTGGCGGACAGA 2887  
2820 GAGAGCAGAGGTTACAGTGAGCAAGATCGTGCCACTGCTTCCAGCTGGCGGACAGA 2879  
2888 GCAAGACTCTGTCTC 2902

|||||  
2880 GCAAGACTCTGTCTC 2894

## RESULT 8

ADB62843

ID ADB62843 standard; cDNA; 2759 BP.

XX AC ADB62843;

XX DT 04-DEC-2003 (first entry)

XX Human cDNA encoding clone OCBBF20111370.

Human; ss; gene; pharmaceutical; diagnostic; gene therapy;  
tissue regeneration; cell regeneration; membrane protein;  
signal transduction-related protein; transcription-related protein;  
osteoporosis; neurological disease; cancer; tumour.

XX Homo sapiens.

Key Location/Qualifiers  
CDS 296..1345  
/\*tag= a

FT FT /product= "Clone OCBBF20111370 protein"

XX PN EP1308459-A2.

XX PD 07-MAY-2003.

XX PF 28-MAR-2002; 2002EP-00007401.

XX PR 05-NOV-2001; 2001JP-00379298.

XX PR 25-JAN-2002; 2002US-00350978.

XX PA (HELI-) HELIX RES INST.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Iehli S;  
XX PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Tamechika I;  
XX PI Seki N, Yoshikawa T, Otsuka M, Negahari K, Masuho Y;

XX DR WPI; 2003-450961/43.

XX DR P-PSDB; ADB64813.

XX DR New polynucleotides and polypeptides, useful for developing a diagnostic  
marker or medicines for regulation of their expression and activity, or  
as targets of gene therapy.

XX PS Claim 1; Page; 222pp; English.

XX CC The invention discloses a polynucleotide comprising a sequence selected  
from 1970 fully defined nucleotide sequences which encode novel  
polypeptides. Also claimed is a polypeptide encoded by the polynucleotide  
or its partial peptide, an antibody binding to the polypeptide or peptide  
of the polynucleotide, immunologically assaying the polypeptide or  
peptide of the polynucleotide by contacting the polypeptide or peptide  
with the antibody of the encoded protein, and observing the binding  
between the two, a transformant carrying the polynucleotide in an  
expressible manner and an antisense polynucleotide. The oligonucleotide  
is useful as a primer for synthesising the polynucleotide, or as a probe  
for detecting the polynucleotide. The polynucleotides and encoded  
proteins are useful as pharmaceutical agents and many disease-related  
genes may be included in them, for developing a diagnostic marker or  
medicines for regulation of their expression and activity, or as targets  
of gene therapy. The genes are involved in tissue and/or cell  
regeneration. Membrane proteins, signal transduction-related proteins,  
transcription-related proteins, disease-related proteins and genes  
encoding them can be used as indicators for diseases (e.g. osteoporosis,  
neurological diseases, cancer, tumours). The cDNA may be used to regulate  
the activity or expression of the encoded protein to treat diseases. The  
sequence presented is a cDNA of the invention. Note: Some of the sequence  
data for this patent is not represented in the printed specification, but

CC is based on sequence information supplied by the European Patent Office.

XX  
SQ Sequence 2759 BP; 654 A; 821 C; 747 G; 537 T; 0 U; 0 Other;

Query Match 89.0%; Score 2598.2; DB 10; Length 2759;  
Best Local Similarity 95.2%; Pred. No. 0;  
Matches 2756; Conservative 0; Mismatches 3; Indels 136; Gaps 2;

QY	8	CTCTGCCAGCCCTGTAGCTGTGGAAAGAGCAGCTACTCTCGAGGCGAGGCGCGCAGGCGGGC	67
Db	1		
QY	68	GGCGATGTAGAGGGGGCGCAGCCGACGCCCGCTGGGGAGCCACCGCTAAACCTGTGCAC	127
Db	61		
QY	128	CCCAACCCACCCCTGTGCACAAAAAGAGCTGGCGGGGCGTGGCCACGCTGGCCCTGGGCTGACCTT	187
Db	121		
QY	188	CCTCGGATGTAGAAATCGCGCCCTGTGGAGGATCTCTTCTCTCTAGGCTCTGAAGGCCCGG	247
Db	129	-----	
QY	248	GGAGCGTGTAGCGATGCCAGCTGCACCCGCGGAGGGCTCGCCTCTGTTTGGCCAGTAAGGA	307
Db	171		
QY	308	GGAGAGGCTGTCTCAGCTGTGAGAGGGGTCAATCCCTGTCTCAAGCCAGTGCCTCTTCCACG	367
Db	231	GGAGAGGCTGTCTCAGCTGTGAGAGGGGTCAATCCCTGTCTTCAAGCCAGTGCCTCTTCCACG	290
QY	368	CTCCCATGGGGACCAACCGAGCCACGCTCCGATGTGAAACGTTGGAGCTGTGAAGAGGAAT	427
Db	291	CTCCCATGGGGACCAACCGAGCCACGCTCCGATGTGAAACGTTGGAGCTGTGAAGAGGAAT	350
QY	428	GGCAGGACGAAAGATCTTCCGAGGCCACTCCCAAGAGAGACGGGGTGGAACTGCTTTGGCA	487
Db	351	GGC-----	
QY	488	GCCCGGTGGAAGACATCTCTCTCTCCCAACACGCTAAATTTCAACGGAGCGCATCGTA	547
Db	354	-----	
QY	548	AGAGGAAGACGCTGTGGGCCCAAGAGATCAAGATTTCTGTGATCAGAGTGAAGGGTCCC	607
Db	405	AGAGGAAGACGCTGTGGGCCCAAGAGATCAAGATTTCTGTGATCAGAGTGAAGGGTCCC	464
QY	608	TGCTGTCCGATGACTTTCTGGATACCCCTGATGACCTGGATATTAACGTGGATGACATCG	667
Db	465	TGCTGTCCGATGACTTTCTGGATACCCCTGATGACCTGGATATTAACGTGGATGACATCG	524
QY	668	AGACCCCGATGTAGACCGACTCGCTGTGGAAATGGCAACGAACTGGAGTGGG	727
Db	525	AGACCCCGATGTAGACCGACTCGCTGTGGAAATGGCAACGAACTGGAGTGGG	584
QY	728	AAGACGACACCCCGTGGCCACCGCCACAGAACATGCCCGGGGACAGCGGATCTATTTCG	787
Db	585	AAGACGACACCCCGTGGCCACCGCCACAGAACATGCCCGGGGACAGCGGATCTATTTCG	644
QY	788	GGGACGGCACGACGGAGGACGGCAGCGCGCCCAACGGGGCCCTGTGGCGGACAGTGTATCA	847
Db	645	GGGACGGCACGACGGAGGACGGCAGCGCGCCCAACGGGGCCCTGTGGCGGACAGTGTATCA	704
QY	848	TGGGGAGCAGAGGCAACCGTATAGACTGTGACATGATTCGGGCTTACATGAAAGTGGTCA	907
Db	705	TGGGGAGCAGAGGCAACCGTATAGACTGTGACATGATTCGGGCTTACATGAAAGTGGTCA	764
QY	908	CCACGAGGGTACTACGGGAAGGGCTCAAGCCCATCATCGTCTTCGACAGCTGCTCC	967
Db	765	CCACGAGGGTACTACGGGAAGGGCTCAAGCCCATCATCGTCTTCGACAGCTGCTGCTCC	824
QY	968	TTCCAGACAGCAGCCTCCCGCAGCTACCACTACATCATGTGAGAACCTCTTCTGTGATCGTCA	1027

```
Db 1905 CATGGTTTTCATTGACCGCGCCCTTCTCGCTCATATGACACCCAGCTCTTTGAGA 1964
Qy 2108 GGATCAGAGCCATTGCAAGAAGAGCCGCTCCCAACCATCTTGTCTCTCCGATTGCAA 2167
Db 1965 GGATCAGAGCCATTGCAAGAAGAGCCGCTCCCAACCATCTTGTCTCTCCGATTGCAA 2024
Qy 2168 AATGACACCCAGTAACTTAGAATCTCAAGCCCTTTAACTCAAGTGTCAAGCCACC 2227
Db 2025 AATGACACCCAGTAACTTAGAATCTCAAGCCCTTTAACTCAAGTGTCAAGCCACC 2084
Qy 2228 GGCAAAACCCGTCATCTCCACCAAGGATGAGATATGAGACCTCACTGCTCCCC 2287
Db 2085 GGCAAAACCCGTCATCTCCACCAAGGATGAGATATGAGACCTCACTGCTCCCC 2144
Qy 2288 CAACCCAGGCTCAGGCTGGGACACGCGCTGTTCCGGGTTGGAAACAGCAGAGCTCA 2347
Db 2145 CAACCCAGGCTCAGGCTGGGACATGCGACGCTGTTCCGGGTTGGAAACAGCAGAGCTCA 2204
Qy 2348 GAAACTGGCTCTGAATATGACAGCTTAGCAAGAGAAAGATACAGGGTATCGGGGTTTG 2407
Db 2205 GAAACTGGCTCTGAATATGACAGCTTAGCAAGAGAAAGATACAGGGTATCGGGGTTTG 2264
Qy 2408 AGTGTTCAGAGCTCAATCGGGAAGATTAATCCAGTGCCTGGCCGACCCACCTGCATT 2467
Db 2265 AGTGTTCAGAGCTCAATCGGGAAGATTAATCCAGTGCCTGGCCGACCCACCTGCATT 2324
Qy 2468 CAAAGCTTGACACGCGGGTCTTGTTCGGGAGGCAAAATTCCTAGGAAAGAAAGACA 2527
Db 2325 CAAAGCTTGACACGCGGGTCTTGTTCGGGAGGCAAAATTCCTAGGAAAGAAAGACA 2384
Qy 2528 GACTTTTCTAATGTGTCAAATGCGGATCACTGTGTCAGATGAGTCTAGAAACACTGAG 2587
Db 2385 GACTTTTCTAATGTGTCAAATGCGGATCACTGTGTCAGATGAGTCTAGAAACACTGAG 2444
Qy 2588 CTCCTGTCTCTGGAAGTATTTAAGAAAGGCTGGGCGGACGATGCTCAGCCTGT 2647
Db 2445 CTCCTGTCTCTGGAAGTATTTAAGAAAGGCTGGGCGGACGATGCTCAGCCTGT 2504
Qy 2648 AATCCAGACTTTGGGAGCGGAGGCGGATCACCTGAGTGAGGAGTTTGAAACA 2707
Db 2505 AATCCAGACTTTGGGAGCGGAGGCGGATCACCTGAGTGAGGAGTTTGAAACA 2564
Qy 2708 GCTGCGCAACATGTTGAAACCTCATCTCTACTTAAATAATACAAATAATGACGAGCGTG 2767
Db 2565 GCTGCGCAACATGTTGAAACCTCATCTCTACTTAAATAATACAAATAATGACGAGCGTG 2624
Qy 2768 TGGCAGGTGCTGTAATCCAGTACTTGGGAGGCTGAGGCGATGAGAACTTAAACCT 2827
Db 2625 TGGCAGGTGCTGTAATCCAGTACTTGGGAGGCTGAGGCGATGAGAACTTAAACCT 2684
Qy 2828 GAGAGGAGAGGTTACAGTGAGCAAGATCGTGCCACTGCTATTCAGGCTGGGCGACAGA 2887
Db 2685 GAGAGGAGAGGTTACAGTGAGCAAGATCGTGCCACTGCTATTCAGGCTGGGCGACAGA 2744
Qy 2888 GCAAGACTGTCTC 2902
Db 2745 GCAAGACTGTCTC 2759
```

RESULT 9  
ADQ63969

ID ADQ63969 standard; cDNA; 2786 BP.

XX AC

XX ADQ63969;

XX XX

DT 07-OCT-2004 (first entry)

XX XX

DE Novel human cDNA sequence #1130.

XX ss;

ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;  
cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;  
neurological disease; Alzheimer's disease; Parkinson's disease; dementia;  
cancer.

```
XX Homo sapiens.
OS
XX EP1440981-A2.
PN
XX 28-JUL-2004.
PD
XX 21-JAN-2004; 2004EP-00001196.
XX 21-JAN-2003; 2003JP-00102206.
PR 09-MAY-2003; 2003JP-00131392.
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;
XX WPI; 2004-535376/52.
DR P-PSDB; ADQ66157.
XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX Claim 1; SEQ ID NO 1130; 2449pp; English.
XX The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a nucleotide
CC sequence of the invention.
XX Sequence 2786 BP; 698 A; 795 C; 742 G; 551 T; 0 U; 0 Other;
Qy Query Match 80.8%; Score 2357.6; DB 12; Length 2786;
XX Best Local Similarity 95.2%; Pred. No. 0;
XX Matches 2492; Conservative 0; Mismatches 4; Indels 122; Gaps 1;
Qy 300 AGTAAGGAGGAGGCTGTCTCAGCTCAGAGGGGTCTATCCCTGCTTCAAGCAGTGCT 359
Db 1 AGTAAGGAGGAGGCTGTCTCAGCTCAGAGGGGTCTATCCCTGCTTCAAGCAGTGCT 60
Qy 360 CTTCCAGCTCCCATGGGACCAAGAGCCAGCTCCGATGGAAGAGCGTGA 419
Db 61 CTTCCAGCTCCCATGGGACCAAGAGCCAGCTCCGATGGAAGAGCGTGA 120
Qy 420 GGAGGAATGGCAGGACGAAGATCTTCCAGGCGCACTCCAGAGAGAGCGGGGTGAACT 479
Db 121 GGAGGAATGGCAGGACGAAGATCTTCCAGGCGCACTCCAGAGAGAGCGGGGTGAACT 180
Qy 480 GCTTGGCAGCGCGTGGGAAGACACATCTCTCTCCCAACCGCTTAAATTTCAACGGAGC 539
Db 181 GCTTGGCAGCGCGTGGGAAGACACATCTCTCTCCCAACCGCTTAAATTTCAACGGAGC 240
Qy 540 GCATCGTAAGGAGAGAGCGTGGTGGCCCGCAGAGATCAACATTTCTCTGGATCAGAGTGA 599
Db 241 GCATCGTAAGGAGAGAGCGTGGTGGCCCGCAGAGATCAACATTTCTCTGGATCAGAGTGA 300
Qy 600 GGGGTCTCTGTCTGCTGATGACTTCTTGGATACCCCTGATGACCTGGATATTAAACGTGA 659
Db 301 GGGGTCTCTGTCTGCTGATGACTTCTTGGATACCCCTG----- 338
Qy 660 TGACATCAGACCCCGGATGAGACCGACTCGTGGAGTTCCTGGGGGATGGCAACGAACT 719
Db 339 ----- 338
Qy 720 GGAGTGGGAAGACGACACCCCGTGGCCACCGCCAGAAACATGCCCGGGGACGCGCGGA 779
Db 339 -----ATGCCCGGGGACGAGCGCGGA 358
```

Qy	780	TCTATTTGGGACGGCACGACGAGAGGACGGCAGCGCGCGCCAAACGGGGCGCCTGTGGCGGAC	839
Db	359	TCTATTTGGGACGGCACGACGAGAGGACGGCAGCGCGCGCGCCAAACGGGGCGCCTGTGGCGGAC	418
Qy	840	AGTGATCATCGGGGAGCAAGAGCACCGPATAGACTGACATGATCGGGCCCTTACATGAA	899
Db	419	AGTGATCATCGGGGAGCAAGAGCACCGPATAGACTGACATGATCGGGCCCTTACATGAA	478
Qy	900	AGTGGTCAACCAACGAGGGGTACTACGGCGAAGGGCTCAAAGCCCATCATCGTCTTCGCAGC	959
Db	479	AGTGGTCAACCAACGAGGGGTACTACGGCGAAGGGCTCAAAGCCCATCATCGTCTTCGCAGC	538
Qy	960	CTGCTTCTTTCAGACAGCAGCCTCCCGCATCACTACATCATGAGAGAACCTTTCCT	1019
Db	539	CTGCTTCTTTCAGACAGCAGCCTCCCGCATCACTACATCATGAGAGAACCTTTCCT	598
Qy	1020	GTACGTCAATCAGCAGCTTAGAGCTCTTGCTGGCTGAGGACTACATGATCGTGTACCTGAA	1079
Db	599	GTACGTCAATCAGCAGCTTAGAGCTCTTGCTGGCTGAGGACTACATGATCGTGTACCTGAA	658
Qy	1080	CGGTGCCACGCGCCCGCGGAGATGCGCTGGAATCGGCTGGCTGAAGAAGTGTCTACCAAGT	1139
Db	659	CGGTGCCACGCGCCCGCGGAGATGCGCTGGAATCGGCTGGCTGAAGAAGTGTCTACCAAGT	718
Qy	1140	GATCGACCGGAGGTTGCGGAAAAAACCCTGAAGTCTCTTGATCATCGTCCACCCCTCGTGTT	1199
Db	719	GATCGGCGGAGGTTGCGGAAAAAACCCTGAAGTCTCTTGATCATCGTCCACCCCTCGTGTT	778
Qy	1200	CATTGGGACTGTGCTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTCAACAAGAT	1259
Db	779	CATTGGGACTGTGCTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTCAACAAGAT	838
Qy	1260	CCAGTACGTGCACAGCTTGGAGGACCTGGAGCAACTCATCCCTATGGAACACGTCACAGAT	1319
Db	839	CCAGTACGTGCACAGCTTGGAGGACCTGGAGCAACTCATCCCTATGGAACACGTCACAGAT	898
Qy	1320	CCCAGACTGCTCTGCAATACGAAGGAAAGACTGAAGGCCAGGAGGAGGAGCGCGAG	1379
Db	899	CCCAGACTGCTCTGCAATACGAAGGAAAGACTGAAGGCCAGGAGGAGGAGCGCGAG	958
Qy	1380	GCCCCAGCGGAGTTGTGCTGCCAGGTCTGAAGAGAGCCAGAGGTGGCACCAAGTGA	1439
Db	959	GCCCCAGCGGAGTTGTGCTGCCAGGTCTGAAGAGAGCCAGAGGTGGCACCAAGTGA	1018
Qy	1440	AAACAGGCTCTGCTCTGAGTCTCAGAAAGATCAGAAACCAAGCATGCTCTGAGGCGACGTCAG	1499
Db	1019	AAACAGGCTCTGCTCTGAGTCTCAGAAAGATCAGAAACCAAGCATGCTCTGAGGCGACGTCAG	1078
Qy	1500	CATTAACAAAGCATGGAAGAGATTCAGATGCGAGAAAACCTCTGTCTAGAGCGCCCACT	1559
Db	1079	CATTAACAAAGCATGGAAGAGATTCAGATGCGAGAAAACCTCTGTCTAGAGCGCCCACT	1138
Qy	1560	GGCCCCAGATCTCATCTGCTCTCATCTGAGTCCCAATCTTCCAGGGTGCCAGGCCCTC	1619
Db	1139	GGCCCCAGATCTCATCTGCTCTCATCTGAGTCCCAATCTTCCAGGGTGCCAGGCCCTC	1198
Qy	1620	GGTTTCACTCTGGAACCCAGCATCTCTTTTTCAGCTGCTTGAACAATGTTATTTTTTTTTT	1679
Db	1199	GGTTTCACTCTGGAACCCAGCATCTCTTTTTCAGCTGCTTGAACAATGTTATTTTTTTTTT	1258
Qy	1680	TTAAACGATGCAATTTTGTGCGTTTCCAGAAAAGGGCCCACTCTGTAGGCCCTCACCCCTC	1739
Db	1259	TTAAACGATGCAATTTTGTGCGTTTCCAGAAAAGGGCCCACTCTGTAGGCCCTCACCCCTC	1318
Qy	1740	CACACTCAGAACTCTCAGCGGAGGAAGGCAAGCGCAGGGGGTGCCCCGGTGCGGT	1799
Db	1319	CACACTCAGAACTCTCAGCGGAGGAAGGCAAGCGCAGGGGGTGCCCCGGTGCGGT	1378
Qy	1800	CGGTGGCTCTCGCTCTCTGCTCGCAGCCCTCTGCTGAGCTGATACAGATTCAAGAC	1859
Db	1379	CGGTGGCTCTCGCTCTCTGCTCGCAGCCCTCTGCTGAGCTGATACAGATTCAAGAC	1438
Qy	1860	CCTTCTCTTGTCTGTCAACCGCTCCAGGTTGGAGGCCACAGACACCAACCGCCACCGCGC	1919

**RESULT 10**



ADRL6283  
ID ADR16283 standard; DNA; 47219 BP.  
XX AC  
XX ADR16283;  
XX DT  
XX 21-OCT-2004 (first entry)  
XX DE Human Cayman ataxia splice site mutant DNA.  
XX KW Human; jittery; ds; Cayman ataxia; ATCAY; KIAA1872; chromosome 19p13.3;  
XX KW ataxia; myoclonus; dystonia; epilepsy; nystagmus; splice site mutant;  
XX KW SNP; single nucleotide polymorphism.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
XX FT misc\_feature 33181..33190  
XX FT /\*tag= a  
XX FT /note= "Exon 9-intron 9 boundary"  
XX FT variation replace(33187,G)  
XX FT /\*tag= b  
XX FT /standard\_name= "Single nucleotide polymorphism"  
XX FT /note= "Causes splice site mutation"  
XX FN US2004146900-A1.  
XX XX  
XX PD 29-JUL-2004.  
XX XX  
XX PP 03-NOV-2003; 2003US-00699941.  
XX XX  
XX PR 01-NOV-2002; 2002US-0422971P.  
XX PR 08-NOV-2002; 2002US-0424973P.  
XX XX  
XX PA (UNMI ) UNIV MICHIGAN.  
XX PI Burmeister M;  
XX XX  
XX DR WPI; 2004-552665/53.  
XX XX  
XX PT Detecting variant Cayman ataxia polypeptide or nucleic acid sequence in  
XX PT subject by detecting presence or absence of variant Cayman ataxia  
XX PT polypeptide or nucleic acid in biological sample.  
XX XX  
XX PS Claim 5; SEQ ID NO 10; 98pp; English.  
XX XX  
XX CC The invention relates to detecting variant Cayman ataxia polypeptide or  
XX CC nucleic acid sequence (ATCAY) in a subject, involving providing a  
XX CC biological sample from a subject, where the biological sample comprises a  
XX CC Cayman ataxia polypeptide or nucleic acid, and detecting the presence or  
XX CC absence of a variant Cayman ataxia polypeptide or nucleic acid in the  
XX CC biological sample. Also included is a kit comprising a reagent for  
XX CC detecting the presence or absence of a variant Cayman ataxia nucleic acid  
XX CC or polypeptide in a biological sample. The human ATCAY gene (previously  
XX CC isolated as cDNA KIAA1872, located on chromosome 19p13.3) was identified  
XX CC and mapped using the fact that it is the homologue of the mouse Jittery  
XX CC gene, located on mouse chromosome 10. The method is useful for detecting  
XX CC variant Cayman ataxia polypeptide or nucleic acid sequence in a subject.  
XX CC The presence of the variant Cayman ataxia polypeptide or nucleic acid is  
XX CC indicative of a disorder chosen from ataxia, polypeptide or nucleic acid  
XX CC epilepsy, and nystagmus in the subject. The biological sample is chosen  
XX CC from blood sample, a tissue sample, urine sample, saliva sample, and an  
XX CC amniotic fluid sample. The subject is chosen from embryo, foetus, newborn  
XX CC animal, young animal, and an adult animal. The animal is a human. The  
XX CC human is an adult female of child-bearing age. The present sequence is  
XX CC the DNA representing a Human ATCAY splice site mutant (disease causing)  
XX CC where a single nucleotide polymorphism in the splice site of exon 9-  
XX CC intron 9 causes a read through mutant protein to be produced.  
XX SQ Sequence 47219 BP; 11562 A; 12203 C; 12045 G; 11209 T; 0 U; 200 Other;  
Query Match 49.3%; Score 1438.2; DB 13; Length 47219;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1440; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1475	CAAGCATGTCTCTGAGCGCAGTGCAGATACAAAGAGCATGGAAGATTCAGATGCC	1534
DB	43720	CTAGCATGTCTCTGAGCGCGAGTGCAGATACAAAGAGCATGGAAGATTCAGATGCC	43779
QY	1535	AGAAAACCTCTGTCTAGAGCCCACTGGCCCCAGATCTCATCTGCTCATCTCTGAGTCCC	1594
DB	43780	AGAAAACCTCTGTCTAGAGCCCACTGGCCCCAGATCTCATCTGCTCATCTCTGAGTCCC	43839
QY	1595	ATCTTCCAAAGGGTCCAGCCCTCCGTTCTCTCTGAAACCCAGCATCTCTTTCAGCTG	1654
DB	43840	ATCTTCCAAAGGGTCCAGCCCTCCGTTCTCTCTGAAACCCAGCATCTCTTTCAGCTG	43899
QY	1655	CTTGAACCATTTGTTATTTTTTTTTTAAACGATGAGTATTTGTGCTTCCAGAAAGGG	1714
DB	43900	CTTGAACCATTTGTTATTTTTTTTTTAAACGATGAGTATTTGTGCTTCCAGAAAGGG	43959
QY	1715	CCAGCTCTGAGCCCTCTCACCTTCCACACTCCAGAACTCTCAGCCGAGGAAGCAAGAA	1774
DB	43960	CCAGCTCTGAGCCCTCTCACCTTCCACACTCCAGAACTCTCAGCCGAGGAAGCAAGAA	44019
QY	1775	GCGAGGGGGTGGCCCGGTGGCGTGGCTCCGCTCCGCTCCGCTCCGAGCCCTTGTGT	1834
DB	44020	GCGAGGGGGTGGCCCGGTGGCGTGGCTCCGCTCCGCTCCGCTCCGAGCCCTTGTGT	44079
QY	1835	CAGAGCTGGATACAAAGATTCAAGACCTTCTCTGCTGTGTACCCCGCTCCAGGTTGGAGC	1894
DB	44080	CAGAGCTGGATACAAAGATTCAAGACCTTCTCTGCTGTGTACCCCGCTCCAGGTTGGAGC	44139
QY	1895	CACAGACACCCACCGCCACCGCGTGGGTCTGCTCTCTCTCTGCTCTCTTCCCTCCAG	1954
DB	44140	CACAGACACCCACCGCCACCGCGTGGGTCTGCTCTCTCTCTGCTCTCTTCCCTCCAG	44199
QY	1955	AATGGGCTCAGACTAGAGCTCAACCCCGCTATGAGGGCCAGCTCTGGGGTAGCTC	2014
DB	44200	AATGGGCTCAGACTAGAGCTCAACCCCGCTATGAGGGCCAGCTCTGGGGTAGCTC	44259
QY	2015	CTGACCTCGACCTTATGTCCAAATTTTCACACCCATGTTTTTCATTTGACCCGCCCT	2074
DB	44260	CTGACCTCGACCTTATGTCCAAATTTTCACACCCATGTTTTTCATTTGACCCGCCCT	44319
QY	2075	TCTCGCTCATATATGACACCCAGCTCTCTTGTAGAGATCAGAGCCCATTTGCAAGAAG	2134
DB	44320	TCTCGCTCATATATGACACCCAGCTCTCTTGTAGAGATCAGAGCCCATTTGCAAGAAG	44379
QY	2135	CCGCTGCCAACCATCTTGTCTCGATGCGAATATGACCCAGTAACTTAGACATT	2194
DB	44380	CCGCTGCCAACCATCTTGTCTCGATGCGAATATGACCCAGTAACTTAGACATT	44439
QY	2195	CTCAAGCCCTTTAACTCAGATGTCAAGCCACCGGGCAACCCCGTCAATACCTCCACC	2254
DB	44440	CTCAAGCCCTTTAACTCAGATGTCAAGCCACCGGGCAACCCCGTCAATACCTCCACC	44499
QY	2255	AAGGAATGAGATATGTGACCTCATCTGTCTCCCAACCCAGCGTCAAGCTGGGACAGCC	2314
DB	44500	AAGGAATGAGATATGTGACCTCATCTGTCTCCCAACCCAGCGTCAAGCTGGGACAGCC	44559
QY	2315	PACGCTGTTCGGGTTGGAAACAGCAGAGGCTCAGAAATGCGCTCTGAAATAGGAGACCT	2374
DB	44560	PACGCTGTTCGGGTTGGAAACAGCAGAGGCTCAGAAATGCGCTCTGAAATAGGAGACCT	44619
QY	2375	AGCAAGAGGAGATACAGGGTATCGGGGCTTTGAGTGTTCAGAAAGTCAATTCGGGAAGAT	2434
DB	44620	AGCAAGAGGAGATACAGGGTATCGGGGCTTTGAGTGTTCAGAAAGTCAATTCGGGAAGAT	44679
QY	2435	AAATCCAGTGGCTGGCCGACCCACCTGCATTCAAAGCTTGGACAGCGGTTCTTGT	2494
DB	44680	AAATCCAGTGGCTGGCCGACCCACCTGCATTCAAAGCTTGGACAGCGGTTCTTGT	44739
QY	2495	CGGGAGGCAAAATTTCCCTAGGAAAGAGACAGACTTTCTTAATGTGTGCCAAATGCG	2554
DB	44740	CGGGAGGCAAAATTTCCCTAGGAAAGAGACAGACTTTCTTAATGTGTGCCAAATGCG	44799



```
QY 2555 ATCACTGGTCAGATGGACTCTAGACGACTGAGCTCCCTGTCTCTCGAAGTATTAAAGAA 2614
Db |||||
QY 44800 ATCACTGGTCAGATGGACTCTAGACGACTGAGCTCCCTGTCTCTCGAAGTATTAAAGAA 44859
Db |||||
QY 2615 AAGGCTGGGCGCAGCAGATGGCTCAGCGCTGTAATCCAGACTTTGGGAGGCCGAGGCA 2674
Db |||||
QY 44860 AAGGCTGGGCGCAGCAGATGGCTCAGCGCTGTAATCCAGACTTTGGGAGGCCGAGGCA 44919
Db |||||
QY 2675 GGCGGATCACTCAGGTGAGGAGTTTGAGAACAGCCTGGCCCAACATGGTGAAACCTTCATC 2734
Db |||||
QY 44920 GGCGGATCACTCAGGTGAGGAGTTTGAGAACAGCCTGGCCCAACATGGTGAAACCTTCATC 44979
Db |||||
QY 2735 TCTACTAATAAATACAAAATATCCAGGCTGGTGGCAGTGCCTGTATCCAGTACT 2794
Db |||||
QY 44980 TCTACTAATAAATACAAAATATCCAGGCTGGTGGCAGTGCCTGTATCCAGTACT 45039
Db |||||
QY 2795 TGGGAGGCTGAGGCAATGAGATCACTTAAACCTGAGAGGCGAGAGTTACAGTGAGCAG 2854
Db |||||
QY 45040 TGGGAGGCTGAGGCAATGAGATCACTTAAACCTGAGAGGCGAGAGTTACAGTGAGCAG 45099
Db |||||
QY 2855 ATCGTCCCACTGCATTCAGCTTCCAGCTTGGGCGACAGAGCAAGACTCTGTCTCAAAAAA 2914
Db |||||
QY 45100 ATCGTCCCACTGCATTCAGCTTCCAGCTTGGGCGACAGAGCAAGACTCTGTCTCAAAAAA 45159
Db |||||
QY 2915 AAA 2917
Db |||||
QY 45160 AAA 45162
Db |||||

RESULT 11
ADRL6284
ID ADRL6284 standard; DNA; 59884 BP.
XX
AC ADRL6284;
XX
DT 21-OCT-2004 (first entry)
XX
DE Human Cayman ataxia genomic DNA.
XX
KW Human; jittery; ds; Cayman ataxia; ATCAY; KIAA1872; chromosome 19p13.3;
KW ataxia; myoclonus; dystonia; epilepsy; nystagmus; splice site mutant;
KW SNP; single nucleotide polymorphism; gene.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_feature 45847..45852
FT /*tag= a
FT /note= "Exon 9-intron 9 boundary"
FT variation replace(45850,T)
FT /*tag= b
FT /standard_name= "single nucleotide polymorphism"
FT /note= "Causes splice site mutation"
XX
US2004146900-A1.
PN
XX
XX
PD 29-JUL-2004.
XX
XX
PP 03-NOV-2003; 2003US-00699941.
XX
XX
PP 01-NOV-2002; 2002US-0422971P.
XX
PP 08-NOV-2002; 2002US-0424973P.
XX
PA (UNMI ) UNIV MICHIGAN.
XX
PI Burmeister M;
XX
XX
DR WPI; 2004-552665/53.
XX
XX
PT Detecting variant Cayman ataxia polypeptide or nucleic acid sequence in
PT subject, by detecting presence or absence of variant Cayman ataxia
PT polypeptide or nucleic acid in biological sample.
XX
```

```
PS
XX
CC The invention relates to detecting variant Cayman ataxia polypeptide or
CC nucleic acid sequence (ATCAY) in a subject, involving providing a
CC biological sample from a subject, where the biological sample comprises a
CC Cayman ataxia polypeptide or nucleic acid, and detecting the presence or
CC absence of a variant Cayman ataxia polypeptide or nucleic acid in the
CC biological sample. Also included is a kit comprising a reagent for
CC detecting the presence or absence of a variant Cayman ataxia nucleic acid
CC or polypeptide in a biological sample. The human ATCAY gene (previously
CC isolated as cDNA KIAA1872, located on chromosome 19p13.3) was identified
CC and mapped using the fact that it is the homologue of the mouse Jittery
CC gene, located on mouse chromosome 10. The method is useful for detecting
CC variant Cayman ataxia polypeptide or nucleic acid sequence in a subject.
CC The presence of the variant Cayman ataxia polypeptide or nucleic acid is
CC indicative of a disorder chosen from ataxia, myoclonus, dystonia,
CC epilepsy, and nystagmus in the subject. The biological sample is chosen
CC from blood sample, a tissue sample, urine sample, saliva sample, and an
CC amniotic fluid sample. The subject is chosen from embryo, foetus, newborn
CC animal, young animal, and an adult animal. The animal is a human. The
CC human is an adult female of child-bearing age. The present sequence is
CC the human genomic DNA for ATCAY.
XX
SQ Sequence 59884 BP; 14656 A; 15794 C; 15192 G; 14042 T; 0 U; 200 Other;
Query Match 49.3%; Score 1438.2; DB 13; Length 59884;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1475 CAAGCATGCTCTGAGGCGACGTGAGCATACAAAGAGACATGGAAGATTCAGATGCC 1534
Db |||||
QY 56385 CTAGCATGCTCTGAGGCGACGTGAGCATACAAAGAGACATGGAAGATTCAGATGCC 56444
Db |||||
QY 1535 AGAAAACCTCTGTCTGAGCGCCCATCTGGCCCGAGATCTCATCTGCTCATCTGAGTCCC 1594
Db |||||
QY 56445 AGAAAACCTCTGTCTGAGCGCCCATCTGGCCCGAGATCTCATCTGCTCATCTGAGTCCC 56504
Db |||||
QY 1595 AATCTTCCAAGGGTCCAGCCCTCCGTTCTATCTCTGAAACCCAGCATCTCTTCAGCTG 1654
Db |||||
QY 56505 AATCTTCCAAGGGTCCAGCCCTCCGTTCTATCTCTGAAACCCAGCATCTCTTCAGCTG 56564
Db |||||
QY 1655 CTGAAAAACATTTGATTTTTTTTTTTTAAACGATGAGTATTTGCGTTCCAGAAAAGG 1714
Db |||||
QY 56565 CTGAAAAACATTTGATTTTTTTTTTTTAAACGATGAGTATTTGCGTTCCAGAAAAGG 56624
Db |||||
QY 1715 CCAGCTCTGAGCCCTTCACTCCCTTCCACACTCAGCAAACTCTCAGCCGAGGAAGCAAGAA 1774
Db |||||
QY 56625 CCAGCTCTGAGCCCTTCACTCCCTTCCACACTCAGCAAACTCTCAGCCGAGGAAGCAAGAA 56684
Db |||||
QY 1775 GCGCAGGGGGTGGCCCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1834
Db |||||
QY 56685 GCGCAGGGGGTGGCCCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 56744
Db |||||
QY 1835 CAGAGCTGATACAGATTTCAAGACCTTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1894
Db |||||
QY 56745 CAGAGCTGATACAGATTTCAAGACCTTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 56804
Db |||||
QY 1895 CACAGACACCCACCGCCACCCCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1954
Db |||||
QY 56805 CACAGACACCCACCGCCACCCCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 56864
Db |||||
QY 1955 AATGGGCTCAGACCTAGAGCTCAACCCCTATGAGGCCAGCTCTGGGGTAGCTC 2014
Db |||||
QY 56865 AATGGGCTCAGACCTAGAGCTCAACCCCTATGAGGCCAGCTCTGGGGTAGCTC 56924
Db |||||
QY 2015 CTGACCTCCGACCTTTATGTCCAAATTTCAACCCATGTTTTCATTTGACCCGCCCT 2074
Db |||||
QY 56925 CTGACCTCCGACCTTTATGTCCAAATTTCAACCCATGTTTTCATTTGACCCGCCCT 56984
Db |||||
QY 2075 TCTCGCTCATATGACACCCAGCTCTCTTTTGAGAGGATCAGAGCCCATTTGCAAGAGAG 2134
Db |||||
QY 56985 TCTCGCTCATATGACACCCAGCTCTCTTTGAGAGGATCAGAGCCCATTTGCAAGAGAG 57044
Db |||||
```

```
QY 2135 CCGCTGCCACGATCTCTTCTCCGATTCGAAATGACACCCAGTAACTAGAACATT 2194
    |||
Db 57045 CCCTGCCACCACTCTCTCTCCGATTCGAAATGACACCCAGTAACTAGAACATT 57104
    |||
QY 2195 CTCAGGCCCTTTAACTCAGATGCTCAAGCCACCGGCAACCCCGTCAATACCTCCACC 2254
    |||
Db 57105 CTCAGGCCCTTTAACTCAGATGCTCAAGCCACCGGCAACCCCGTCAATACCTCCACC 57164
    |||
QY 2255 AAGGAATGAGATATGAGACCTCACTGCTCCCAACCCAGCGTCAAGCTGGGACACGCC 2314
    |||
Db 57165 AAGGAATGAGATATGAGACCTCACTGCTCCCAACCCAGCGTCAAGCTGGGACACGCC 57224
    |||
QY 2315 AAGCTGTTCCGGTTGGAAACAGACAGGCTCAGAACTGGCTCTGAAATAGGACACT 2374
    |||
Db 57225 AAGCTGTTCCGGTTGGAAACAGACAGGCTCAGAACTGGCTCTGAAATAGGACACT 57284
    |||
QY 2375 AGCAAGAGGAAGATACAGGGTATCGGGGCTTTCAGTGTTCAGAGTCAATTCGGGAGAT 2434
    |||
Db 57285 AGCAAGAGGAAGATACAGGGTATCGGGGCTTTCAGTGTTCAGAGTCAATTCGGGAGAT 57344
    |||
QY 2435 AAATCCAGTGGCTGGCCGACGACCTGCAATCAAAAGCTTGGACACGGGTTCTTGT 2494
    |||
Db 57345 AAATCCAGTGGCTGGCCGACGACCTGCAATCAAAAGCTTGGACACGGGTTCTTGT 57404
    |||
QY 2495 CGGAGGCAAAATTCCTTAGGAAAGAGACAGACTTTTCTAATGTGTCCAAATGCGG 2554
    |||
Db 57405 CGGAGGCAAAATTCCTTAGGAAAGAGACAGACTTTTCTAATGTGTCCAAATGCGG 57464
    |||
QY 2555 ATCACTGGTCAGATGGACTCTAGACACTGAGCTCCCTGCTCTGGAAGTATTTAAGA 2614
    |||
Db 57465 ATCACTGGTCAGATGGACTCTAGACACTGAGCTCCCTGCTCTGGAAGTATTTAAGA 57524
    |||
QY 2615 AAGGCTGGGCGAGGACGATGGCTCAGCGCTGTAATCCAGACTTTGGGAGGCCGAGGCA 2674
    |||
Db 57525 AAGGCTGGGCGAGGACGATGGCTCAGCGCTGTAATCCAGACTTTGGGAGGCCGAGGCA 57584
    |||
QY 2675 GCGGATACCTGAGTGAGAGTTTGAGAAACAGCTTGCGCAACATGTTGAAACCTCATC 2734
    |||
Db 57585 GCGGATACCTGAGTGAGAGTTTGAGAAACAGCTTGCGCAACATGTTGAAACCTCATC 57644
    |||
QY 2735 TCTACTAAATAACAAATAATAGCAGGCTGCTGAGGCTGCTGTAATCCGCTACT 2794
    |||
Db 57645 TCTACTAAATAACAAATAATAGCAGGCTGCTGAGGCTGCTGTAATCCGCTACT 57704
    |||
QY 2795 TGGAGGCTGAGGATGAGAAATCACTTAAACCTGAGAGGCGAGAGTTACAGTGAGCAG 2854
    |||
Db 57705 TGGAGGCTGAGGATGAGAAATCACTTAAACCTGAGAGGCGAGAGTTACAGTGAGCAG 57764
    |||
QY 2855 ATGCTGCCACTGATTCAGCTGGGCGACAGAGCAAGACTCTGTCTCAAAAAAATAA 2914
    |||
Db 57765 ATGCTGCCACTGATTCAGCTGGGCGACAGAGCAAGACTCTGTCTCAAAAAAATAA 57824
    |||
QY 2915 AAA 2917
    |||
Db 57825 AAA 57827
```

## RESULT 12

AD862832  
ID AD862832 standard; cDNA; 2625 BP.

XX

AC

AD862832;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Key Location/Qualifiers  
CDS 1990..2334  
/\*tag=a  
/product="Clone OCBBF20001780 protein"

EP1308459-A2.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Key Location/Qualifiers  
CDS 1990..2334  
/\*tag=a  
/product="Clone OCBBF20001780 protein"

EP1308459-A2.

07-MAY-2003.

28-MAR-2002; 2002EP-00007401.

05-NOV-2001; 2001JP-00379298.

25-JAN-2002; 2002US-00350978.

(HELI-) HELIX RES INST.

(REAS-) RES ASSOC BIOTECHNOLOGY.

Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

WPI; 2003-450961/43.

P-PSDB; ADB64802.

New polynucleotides and polypeptides, useful for developing a diagnostic marker or medicines for regulation of their expression and activity, or as targets of gene therapy.

Claim 1; Page; 222pp; English.

The invention discloses a polynucleotide comprising a sequence selected from 1970 fully defined nucleotide sequences which encode novel polypeptides. Also claimed is a polypeptide encoded by the polynucleotide or its partial peptide, an antibody binding to the polypeptide or peptide of the polynucleotide, immunologically assaying the polypeptide or peptide of the polynucleotide by contacting the polypeptide or peptide with the antibody of the encoded protein, and observing the binding between the two, a transformant carrying the polynucleotide in an expressible manner and an antisense polynucleotide. The oligonucleotide is useful as a primer for synthesizing the polynucleotide, or as a probe for detecting the polynucleotide. The polynucleotides and encoded proteins are useful as pharmaceutical agents and many disease-related genes may be included in them, for developing a diagnostic marker or medicines for regulation of their expression and activity, or as targets of gene therapy. The genes are involved in tissue and/or cell regeneration. Membrane proteins, signal transduction-related proteins, transcription-related proteins, disease-related proteins and genes encoding them can be used as indicators for diseases (e.g. osteoporosis, neurological diseases, cancer, tumours). The cDNA may be used to regulate the activity or expression of the encoded protein to treat diseases. The sequence presented is a cDNA of the invention. Note: Some of the sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office.

Sequence 2625 BP; 660 A; 754 C; 685 G; 526 T; 0 U; 0 Other;

## Query Match

Best Local Similarity 45.4%; Score 1325.8; DB 10; Length 2625;

Matches 1338; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY	1562	CCCCAGATCTCATCTGCTGCTCATCTGAGTCCCAATCTTCAAGGGTCCAGCCCTCG 1621
Db	1286	CTCCAGATCTCATCTGCTGCTCATCTGAGTCCCAATCTTCAAGGGTCCAGCCCTCG 1345
QY	1622	TTCAATCTGAAACCCAGCATCTTTTTCAGTGTGTGAAACATGTATTTTTTTTTT 1681
Db	1346	TTCAATCTGAAACCCAGCATCTTTTTCAGTGTGTGAAACATGTATTTTTTTTTT 1404
QY	1682	AACGATGAGTATTTTGTGGTTCAGAAAGGGCCAGCTCTGAGCCCTCAGCCCTTCA 1741
Db	1405	AACGATGAGTATTTTGTGGTTCAGAAAGGGCCAGCTCTGAGCCCTCAGCCCTTCA 1464

1742 CACTCAGCACTCTCAGCCGAGGAGGCAAGAGCGCAGGGGGTGGCCCGCGTGGCGTGG 1801  
1465 CACTCAGCACTCTCAGCCGAGGAGGCAAGAGCGCAGGGGGTGGCCCGCGTGGCGTGG 1524  
1802 GTGGCCCTCCGCTCTCTGCTCCAGCCCTCTGCTGGTGGTGGATCAAGATTCACAGACC 1861  
1525 GTGGCCCTCCGCTCTCTGCTCCAGCCCTCTGCTGGTGGTGGATCAAGATTCACAGACC 1584  
1862 TTCTCTGTTGTTCACCCGCTCCAGGTTGGAGCCACAGACACCACCGCCACCCCGGGCTG 1921  
1585 TTCTCTGTTGTTCACCCGCTCCAGGTTGGAGCCACAGACACCACCGCCACCCCGGGCTG 1644  
1922 GGTCTGGTCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1981  
1645 GGTCTGGTCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1704  
1982 CCCCCCTATGAGGCGCCAGCTCTGCTGGGTAGCTCTGCTGCTCTGCTCTGCTCTGCTCT 2041  
1705 CCCCCCTATGAGGCGCCAGCTCTGCTGGGTAGCTCTGCTGCTCTGCTCTGCTCTGCTCT 1764  
2042 CACACCCATGGTTTTCAITTTGACCCGCCCCCTTCTGCTCTGCTCTGCTCTGCTCTGCT 2101  
1765 CACACCCATGGTTTTCAITTTGACCCGCCCCCTTCTGCTCTGCTCTGCTCTGCTCTGCT 1824  
2102 TTGAGAGGATCAGAGCCCATTTGACCAAGAGAGCCGCTGCTCAACCATCTTGTCTCCGA 2161  
1825 TTGAGAGGATCAGAGCCCATTTGACCAAGAGAGCCGCTGCTCAACCATCTTGTCTCCGA 1884  
2162 TTGCAAAATGACACCCAGTATCTAGAACATCTCTCAAGCCCTTTAACTCAGATGTCAA 2221  
1885 TTGCAAAATGACACCCAGTATCTAGAACATCTCTCAAGCCCTTTAACTCAGATGTCAA 1944  
2222 GCCACCGGGCAAAACCCCGCTCAATACCTCCACCAAGGAATGAGATGTTGGACCTCAGTG 2281  
1945 GCCACCGGGCAAAACCCCGCTCAATACCTCCACCAAGGAATGAGATGTTGGACCTCAGTG 2004  
2282 CTCCTCCCAACCCAGAGCTCAGGCTGGGACAGCCCAAGCTGTTCCGGTTGGCAACAGCAGA 2341  
2005 CTCCTCCCAACCCAGAGCTCAGGCTGGGACAGCCCAAGCTGTTCCGGTTGGCAACAGCAGA 2064  
2342 GGTCTCAGAACTGGCTCTGAAATAGGACGACCTAGCAGAGGAGATACAGGGTATCGGG 2401  
2065 GGTCTCAGAACTGGCTCTGAAATAGGACGACCTAGCAGAGGAGATACAGGGTATCGGG 2124  
2402 CGTTTGAGTGTTCAGAACTGTTCCGGAAGATTAATCCAGTCCGCTGGCCCGCAGCCACC 2461  
2125 CGTTTGAGTGTTCAGAACTGTTCCGGAAGATTAATCCAGTCCGCTGGCCCGCAGCCACC 2184  
2462 TGCATTCAAAGCTTGGACAGCGGGTCTTGTTCGGGAGGCAAAATTCCTTAGGAAAAG 2521  
2185 TGCATTCAAAGCTTGGACAGCGGGTCTTGTTCGGGAGGCAAAATTCCTTAGGAAAAG 2244  
2522 AAGACAGCTTTTCTAATGTGTCGAAATGCGGATCAGTGTGATGACCTCTAGAGC 2581  
2245 AAGACAGCTTTTCTAATGTGTCGAAATGCGGATCAGTGTGATGACCTCTAGAGC 2304  
2582 ACTGAGCTCCCTGCTCTCAGAACTTTAAGAAAAGGCTGGCCAGGACGATGCTCAG 2641  
2305 ACTGAGCTCCCTGCTCTCAGAACTTTAAGAAAAGGCTGGCCAGGACGATGCTCAG 2364  
2642 GCTGTAATCCAGACTTTGGAGGCGCAGGAGGCGGATCACTGAGTGAGGAGTTTG 2701  
2365 GCTGTAATCCAGACTTTGGAGGCGCAGGAGGCGGATCACTGAGTGAGGAGTTTG 2424  
2702 AAGACAGCTTTGCGCAATGTTGAACTCTCTCTAATAAATAAATAAATAAATAAATAA 2761  
2425 AAGACAGCTTTGCGCAATGTTGAACTCTCTCTAATAAATAAATAAATAAATAAATAA 2484  
2762 GGTGTGTGAGCTGCTGCTTAATCCAGCTTCTGGAGGCTGAGGATGAGATCACTT 2821  
2485 GGTGTGTGAGCTGCTGCTTAATCCAGCTTCTGGAGGCTGAGGATGAGATCACTT 2544  
2822 AAACCTGAGAGGAGGAGGTTTACAGTGAGCCAGATCGTGCCACTGCACTGCCAGCTGGGC 2881

Db 2545 AAACCTGAGAGGAGGAGGTTTACAGTGAGCCAGATCGTGCCACTGATCCAGCTGGGC 2604  
QY 2882 GACAGAGCAAGACTCTGTCTC 2902  
Db 2605 GACAGAGCAAGACTCTGTCTC 2625

## RESULT 13

ADRL6274  
ID ADR16274 standard; cDNA; 2166 BP.  
XX AC ADR16274;  
XX DT 21-OCT-2004 (first entry)  
XX DE Mouse jittery cDNA.  
KW Mouse; jittery; ss; gene; Cayman ataxia; ATCAY; KIAA1872; chromosome 10;  
XX ataxia; myoclonus; dystonia; epilepsy; nyctagmus.  
OS Mus musculus.  
XX FH Key Location/Qualifiers  
CDS 378..1496  
FT /tag= a  
FT /product= "Jittery"  
FT /transl\_except= (pos:1044..1046,aa:Glu)  
XX PN US2004146900-A1.  
XX PD 29-JUL-2004.  
XX PP 03-NOV-2003; 2003US-00699941.  
PR 01-NOV-2002; 2002US-0422971P.  
PR 08-NOV-2002; 2002US-0424973P.  
XX PA (UNMI ) UNIV MICHIGAN.  
XX PI Burmeister M;  
XX WPI; 2004-552665/53.  
DR P-PSDB; ADR16275.  
XX PT Detecting variant Cayman ataxia polypeptide or nucleic acid sequence in  
PT subject, by detecting presence or absence of variant Cayman ataxia  
PT polypeptide or nucleic acid in biological sample.  
XX PS Disclosure; SEQ ID NO 1; 98pp; English.

CC The invention relates to detecting variant Cayman ataxia polypeptide or  
CC nucleic acid sequence (ATCAY) in a subject, involving providing a  
CC biological sample from a subject, where the biological sample comprises a  
CC Cayman ataxia polypeptide or nucleic acid, and detecting the presence or  
CC absence of a variant Cayman ataxia polypeptide or nucleic acid in the  
CC biological sample. Also included is a kit comprising a reagent for  
CC detecting the presence or absence of a variant Cayman ataxia nucleic acid  
CC or polypeptide in a biological sample. The human ATCAY gene (previously  
CC isolated as cDNA KIAA1872, located on chromosome 19p13.3) was identified  
CC and mapped using the fact that it is the homologue of the mouse Jittery  
CC gene, located on mouse chromosome 10. The method is useful for detecting  
CC variant Cayman ataxia polypeptide or nucleic acid sequence in a subject.  
CC The presence of the variant Cayman ataxia polypeptide or nucleic acid is  
CC indicative of a disorder chosen from ataxia, myoclonus, dystonia,  
CC epilepsy, and nyctagmus in the subject. The biological sample is chosen  
CC from blood sample, a tissue sample, urine sample, saliva sample, and an  
CC amniotic fluid sample. The subject is chosen from embryo, foetus, newborn  
CC animal, young animal, and an adult animal. The animal is a human. The  
CC human is an adult female of child-bearing age. The present sequence is  
CC the cDNA encoding the mouse Jittery protein.

XX SQ Sequence 2166 BP; 523 A; 619 C; 576 G; 448 T; 0 U; 0 Other;



PT CDNAs.

PS Claim 1; SEQ ID NO 5423; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX Sequence 790 BP; 166 A; 245 C; 251 G; 125 T; 0 U; 3 Other;

Query Match 25.5%; Score 745.2; DB 4; Length 790;

Best Local Similarity 99.2%; Pred. No. 1.5e-171;

Matches 758; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 8 CTCTGCAGCCCTGAGCTGGGAAGACAGCTACCTCGAGGCGAGGCGCGAGGCGGC 67

DB 1 CTCTGCAGCCCGGCTGGGAAGACAGCTACCTCGAGGCGAGGCGCGAGGCGGC 60

QY 68 GCGATGAGAGGGGGCGAGCGGAGCCCGCGCTGGGAGCCACCGCTAACCTTGAC 127

DB 61 GCGATGAGAGGGGGCGAGCGGAGCCCGCGCTGGGAGCCACCGCTAACCTTGAC 120

QY 128 CCCACCCCTCGACAAAGAGCTGGCGGGCGCTGGCCAGCTGCGCCCTGGTGACCTT 187

DB 121 CCCACCCCTCGACAAAGAGCTGGCGGGCGCTGGCCAGCTGCGCCCTGGTGACCTT 180

QY 188 CTTGGATGAGAAATCCGCGCTCGAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 247

DB 181 CTTGGATGAGAAATCCGCGCTCGAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240

QY 248 GGAGCGTGAAGATGCCAGCTGCACCCCGGCGAGGCGCTGCGCTTTGTTGCCAGTAAGA 307

DB 241 GGAGCGTGAAGATGCCAGCTGCACCCCGGCGAGGCGCTGCGCTTTGTTGCCAGTAAGA 300

QY 308 GGAGGCTGTCTAGCTCAGAGGGGTATCCCTGCTTCAAGCCAGTGCCTCTTCCAG 367

DB 301 GGAGGCTGTCTAGCTCAGAGGGGTATCCCTGCTTCAAGCCAGTGCCTCTTCCAG 360

QY 368 CTTCCATGGGACACCGAGCCAGCTCCGATGGAACCTGGAACCTGTTGCA 427

DB 361 CTTCCATGGGACACCGAGCCAGCTCCGATGGAACCTGGAACCTGTTGCA 420

QY 428 GGCAGGACGAATCTTCCAGGCCACTCCAGAGAGACGGGGTGGAACTGTTGCA 487

DB 421 GGCAGGACGAATCTTCCAGGCCACTCCAGAGAGACGGGGTGGAACTGTTGCA 480

QY 488 GCGCGTGAAGACATCTCTCTCTCCCAACCGCTAAATTCACCGAGCGCATCGTA 547

DB 481 GCGCGTGAAGACATCTCTCTCTCCCAACCGCTAAATTCACCGAGCGCATCGTA 540

QY 548 AGAGGAGACGCTGGTGGCCCGAGAGATCAACATTTCTTGATCAGATGAGGGTCCC 607

DB 541 AGAGGAGACGCTGGTGGCCCGAGAGATCAACATTTCTTGATCAGATGAGGGTCCC 600

QY 608 TCGTGTCCGATGACATCTTTGGATACCCCTGATGACCTGGATATTAACGTGGATGACATCG 667  
DB 601 TCGTGTCCGATGACATCTTTGGATACCCCTGATGACCTGGATATTAACGTGGATGACATCG 660  
QY 668 AGACCCCGGATGAGACCGGACTCGCTGGAGTTCTTGGGGAATGGCAACACTGGATGGG 727  
DB 661 AGACCCCGGATGAGACCGGACTCGCTGGAGTTCTTGGGGAATGGCAACACTGGATGGG 720  
QY 728 AAGAGCAGACACCCCGGCTGGCCACC-GCCAAAGAACATGCCCGGGGA 770  
DB 721 AAGAGCAGACACCCCGGCTGGCCACC-GCCAAAGAACATGCCCGGGGA 764

RESULT 15

AAS80323

ID AAS80323 standard; cDNA; 703 BP.

XX AAS80323;

AC AAS80323;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #16127.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG16136.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity.

XX Claim 1; SEQ ID NO 16127; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

XX sequences. (I) is useful as hybridisation probes, polymerase chain

XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

XX and in recombinant production of (II). The polynucleotides are also used

XX in diagnostics as expressed sequence tags for identifying expressed

XX genes. (I) is useful in gene therapy techniques to restore normal

XX activity of (II) or to treat disease states involving (II). (II) is

XX useful for generating antibodies against it, detecting or quantitating a

XX polypeptide in tissue, as molecular weight markers and as a food

XX supplement. (II) and its binding partners are useful in medical imaging

XX of sites expressing (II). (I) and (II) are useful for treating disorders

XX involving aberrant protein expression or biological activity. The

XX polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits to assess biodiversity

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic

XX coding sequences of the invention. Note: The sequence data for this

XX patent did not appear in the printed specification, but was obtained in

XX electronic format directly from WIPO at

XX ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 703 BP; 152 A; 218 C; 213 G; 120 T; 0 U; 0 Other;  
Query Match 23.1%; Score 675; DB 5; Length 703;  
Best Local Similarity 98.4%; Pred. No. 2e-154;  
Matches 692; Conservative 0; Mismatches 10; Indels 1; Gaps 1;  
QY 66 GCGGCGATGAGAGGGGCGAGCGCGAGCCCGCGCTGGGAGCCACCGCTAACCCCTGC 125  
DB 1 GCAGAGATGAGAGGGGCGAGGCTCAGCCCGCGCTGGGGAGCCACCGCTAACCCCTGC 60  
QY 126 ACCCCACCCACCCCTGCACAAAAGAGCTGGCGGGCGCTGCCACGTCGCCCTGGGGTGACC 185  
DB 61 ACCCCACCCACCCCTGCACAAAAGAGCTGGCGGGCGCTGGCCACGTCGCCCTGGGGTGACC 120  
QY 186 TTCCTCGGATGCAAGATCGCCCCCTCGAGCATCTCTTCTCCTTAGGCTCTGAAGGCC 245  
DB 121 TTCCTCGGATGCAAGATCGCCCCCTCGAGCATCTCTTCTCCTTAGGCTCTGAAGGCC 180  
QY 246 GGGGAGGTGAGCGATGCCCGAGCTGCACCGGGGAGGGCTGGCTTTGTTGGCA-GTAA 304  
DB 181 GGGGAGGTGAGCGATGCCCGAGCTGCACCGGGGAGGGCTGGCTTTGTTGGCATGTAA 240  
QY 305 GAGGAGAGGCTGTCTCAGCTGCAGAGGGGTCACTCCCTGCTCAAGCCAGTGCCTCTTCC 364  
DB 241 GAGGAGAGGCTGTCTCAGCTGCAGAGGGGTCACTCCCTGCTCAAGCCAGTGCCTCTTCC 300  
QY 365 CAGCTCCCATGGGGACCAACGAGCCAGCTCCGGATGGAAACGTGGACGTGAAGGAG 424  
DB 301 CAGCTCCCATGGGGACCAACGAGCCAGCTCCGGATGGAAACGTGGACGTGAAGGAG 360  
QY 425 AATGGCAGGACGAAGATCTTCCAGGCCACTCCAGAGAGAGCGGGGTGGAACCTGCTTG 484  
DB 361 AATGGCAGGACTAAGATCTTCCAGGCCACTCCAGAGAGAGCGGGGTGGAACCTGCTTG 420  
QY 485 GCAGCCCGTGGAGACACATCTCTCTCCCAACGCTAAATTTCAACGGAGCGCATC 544  
DB 421 GCAGCCCGTGGAGACACATCTCTCTCCCAACGCTAAATTTCAACGGAGCGCATC 480  
QY 545 GTAAGAGGAGACGCTGGTGGCCCCAGAGATCAACATTTCTCTGGATCAGAGTGAGGGGT 604  
DB 481 GCAAGAGGAGACGCTGGTGGCCCCAGAGATCAACATTTCTCTGGATCAGTGAGGGGT 540  
QY 605 CCCTGCTGTCGATGACTTCTTGGATACCCCTGATGACCTGGATATTAACGTGGATGACA 664  
DB 541 CCCTGCTGTCGATGACTTCTTGGATACCCCTGATGACCTGGATATTAACGTGGATGACA 600  
QY 665 TCGAGACCCCGATGAGACCGACTCGCTGGAGTTCTTGGGATGGCAACGAACTGGAGCT 724  
DB 601 TCGAGACCCCGATGAGACCGACTCGCTGGAGTTCTTGGGATGGCAACGAACTGGAGT 660  
QY 725 GGAAGACGACACCCCGTGGCCACCGCCCAAGAACATGCCCGG 767  
DB 661 GGAAGACGACACCCCGTGGCCACCGCCCAAGAACATGCCCGG 703

Search completed: March 5, 2006, 03:30:55  
Job time : 1683 secs



Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	2887.6	99.0	3076	3	US-09-174-937-1	Sequence 1, Appli	
2	2598.2	89.0	2759	3	US-10-104-047-997	Sequence 997, App	
3	1325.8	45.4	2625	3	US-10-104-047-986	Sequence 986, App	
4	1113	38.1	1113	3	US-09-174-937-3	Sequence 3, Appli	
5	309.8	10.6	322	3	US-09-621-976-14512	Sequence 14512, A	
6	302.2	10.4	966	3	US-09-174-937-6	Sequence 6, Appli	
7	302.2	10.4	4235	3	US-09-174-937-4	Sequence 4, Appli	
8	241.8	8.3	103987	3	US-09-949-016-12513	Sequence 12513, A	
9	241.8	8.3	103988	3	US-09-949-016-17050	Sequence 17050, A	
10	241.2	8.3	33712	3	US-09-949-016-15793	Sequence 15793, A	
11	241.2	8.3	193169	3	US-09-949-016-15091	Sequence 15091, A	
12	240.4	8.2	87734	3	US-09-949-016-17521	Sequence 17521, A	
13	239.8	8.2	49378	3	US-09-949-016-13408	Sequence 13408, A	
14	239.6	8.2	11929	3	US-09-949-016-17290	Sequence 17290, A	
15	239.4	8.2	24221	3	US-09-949-016-14964	Sequence 14964, A	
16	239.2	8.2	21590	3	US-09-949-016-13372	Sequence 13372, A	
17	239.2	8.2	21590	3	US-09-949-016-13373	Sequence 13373, A	
18	239	8.2	601	3	US-09-949-016-183967	Sequence 183967, A	
19	239	8.2	601	3	US-09-949-016-184072	Sequence 184072, A	
20	239	8.2	601	3	US-09-949-016-184177	Sequence 184177, A	
21	239	8.2	601	3	US-09-949-016-184282	Sequence 184282, A	
22	239	8.2	43543	3	US-09-778-961-3	Sequence 3, Appli	
23	239	8.2	246230	3	US-09-949-016-17019	Sequence 17019, A	
24	239	8.2	246230	3	US-09-949-016-17020	Sequence 17020, A	

Db	321	CTGCAGAGGGGTCA	TCCCTCTGCTTCAGGCCAGTGCCTTTCCCAAGCTCCCATGGGACAC	380	
Qy	384	CGAAGCCACGCTCCGGATGGAAAA	CGTGGACGTGAAGCAGGAATGGCAGGACGAAGATCT	443	
Db	381	CGAAGCCACGCTCCGGATGGAAAA	ACGTGGACGTGAGCTGAGAGGAGATGGCAGGACGAAGATCT	440	
Qy	444	TCCAGGCCACTCC	CAGAAGAGACGGGGGTGGAACTGCTTTGGCAGCCCGGTGGAAGACAC	503	
Db	441	TCCAGGCCACTCC	CAGAAGAGACGGGGGTGGAACTGCTTTGGCAGCCCGGTGGAAGACAC	500	
Qy	504	ATCCTCTCTCC	CAACACGCTAAATTTCAACGGAGCGCATCGTAAAGGAAAGACGCTGGT	563	
Db	501	ATCCTCTCTCC	CAACACGCTAAATTTCAACGGAGCGCATCGTAAAGGAAAGACGCTGGT	560	
Qy	564	GGCCCCAGAGATCA	CAATTTCTCTCGGATCAGAGTGAGGGGTCCCTGCTGTCGATGACTT	623	
Db	561	GGCCCCAGAGATCA	CAATTTCTCTCGGATCAGAGTGAGGGGTCCCTGCTGTCGATGACTT	620	
Qy	624	CTTTGGATACCC	CTGATGACCTGGATATTAACGTGGATGACATCGAGACCCCGATGAGAC	683	
Db	621	CTTTGGATACCC	CTGATGACCTGGATATTAACGTGGATGACATCGAGACCCCGATGAGAC	680	
Qy	684	CGACTCCCTG	AGTTCCTGGGGAATGGCAACGAACTGGAGTGGGAAGACGACACCCCCGT	743	
Db	681	CGACTCCCTG	AGTTCCTGGGGAATGGCAACGAACTGGAGTGGGAAGACGACACCCCCGT	740	
Qy	744	GGCCACCGCC	CAAGAAATGCTCCGGGGACAGCGGGATCTATTTGGGACGGCACGACGGA	803	
Db	741	GGCCACCGCC	CAAGAAATGCTCCGGGGACAGCGGGATCTATTTGGGACGGCACGACGGA	800	
Qy	804	GGACGGCAGCG	CCGCCAACCGCGGCCCTGTGTGGCGGACAGTGATCATCGGGGAGCAAGACGA	863	
Db	801	GGACGGCAGCG	CCGCCAACCGCGGCCCTGTGTGGCGGACAGTGATCATCGGGGAGCAAGACGA	860	
Qy	864	CCGTATAGACCT	GACATGATTCGGCCCTTACATGAAGTGGTCAACCACGAGGGTACTA	923	
Db	861	CCGTATAGACCT	GACATGATTCGGCCCTTACATGAAGTGGTCAACCACGAGGGTACTA	920	
Qy	924	CGGCGAAGGCT	CAACGCCATCATCGTCTTCGACGCGCTCTTCCTCCAGCAGCAGCCT	983	
Db	921	CGGCGAAGGCT	CAACGCCATCATCGTCTTCGACGCGCTCTTCCTCCAGCAGCAGCCT	980	
Qy	984	CCCCGACTACCA	CTACATCATGGAGAACCTCTTCTCTGTAAGTCAAGCAGCTTAGAGCT	1043	
Db	981	CCCCGACTACCA	CTACATCATGGAGAACCTCTTCTCTGTAAGTCAAGCAGCTTAGAGCT	1040	
Qy	1044	CTTGGTGGCTG	AGGACTACATGATCGTGTACTTGAA	CGGTGCCACGCCCGGGAGGAT	1103
Db	1041	CTTGGTGGCTG	AGGACTACATGATCGTGTACTTGAA	CGGTGCCACGCCCGGGAGGAT	1100
Qy	1104	GCCTGGAA	TGGCTGGTAGAGTGTCTACCATGATCGACCGGAGGTTGCGGAAAA	1163	
Db	1101	GCCTGGAA	TGGCTGGTAGAGTGTCTACCATGATCGACCGGAGGTTGCGGAAAA	1160	
Qy	1164	CCTGAA	GCTCTTGATCATCGTCCACCCCTCGTGGTTCATTCGGACTGTGTGGCCATCTC	1223	
Db	1161	CCTGAA	GCTCTTGATCATCGTCCACCCCTCGTGGTTCATTCGGACTGTGTGGCCATCTC	1220	
Qy	1224	TGCGCCTTT	CAATCAGCTCAAGTTTCATCAACAGATCCAGTACGTCGACAGCTTGGGAAGA	1283	
Db	1221	TGCGCCTTT	CAATCAGCTCAAGTTTCATCAACAGATCCAGTACGTCGACAGCTTGGGAAGA	1280	
Qy	1284	CCTGGAGCA	ACTCATCCCTATGGAAACAGTCCAGATCCAGATCGCTCTGCAATACGA	1343	
Db	1281	CCTGGAGCA	ACTCATCCCTATGGAAACAGTCCAGATCCAGATCGCTCTGCAATACGA	1340	
Qy	1344	AGAGGAAG	AGACTGAAGCCAGGAGAGCGCGAGGCCCCAGCGCGAGTTTGTGCTGCC	1403	
Db	1341	AGAGGAAG	AGACTGAAGCCAGGAGAGCGCGAGGCCCCAGCGCGAGTTTGTGCTGCC	1400	
Qy	1404	CAGGCTG	GAAGAGCCAGGTTGGCACCATGGAAAAACAGGTTCTCTGTGTTCTCAGA	1463	

QY 2544 TCCAAATGCCGATCACTGGTCAATGGAATCTTAGAGCACTGAGCTCCCTGTCTCTGGAA 2603  
DB 2541 TCCAAATGCCGATCACTGGTCAATGGAATCTTAGAGCACTGAGCTCCCTGTCTCTGGAA 2600  
QY 2604 GTATTAAAGAAAAGGCTGGGCCAGGACGATGGCTCAGCCCTGTAATCCAGACTTTGGG 2663  
DB 2601 GTATTAAAGAAAAGGCTGGGCCAGGACGATGGCTCAGCCCTGTAATCCAGACTTTGGG 2660  
QY 2664 AGCCGAGGAGGAGGATCACTGAGTGAAGAGTTTGAAGACAGGCTGGCCAAATGTT 2723  
DB 2661 AGCCGAGGAGGAGGATCACTGAGTGAAGAGTTTGAAGACAGGCTGGCCAAATGTT 2720  
QY 2724 GAAACCTCATCTCTACTTAAATAATCAAAATTAAGCAGGCGTGGTGGCAGGTTGCTGTAA 2783  
DB 2721 GAAACCTCATCTCTACTTAAATAATCAAAATTAAGCAGGCGTGGTGGCAGGTTGCTGTAA 2780  
QY 2784 TCCAGCTACTTGGGAGGCTGAGGCATGAGATCACTTAACTGAGAGGAGGTTTAC 2843  
DB 2781 TCCAGCTACTTGGGAGGCTGAGGCATGAGATCACTTAACTGAGAGGAGGTTTAC 2840  
QY 2844 AGTGAGCAGAGTCTGTCCTCACTGATTCAGGCTGGCGACAGCAAGACTCTGTCTCA 2903  
DB 2841 AGTGAGCAGAGTCTGTCCTCACTGATTCAGGCTGGCGACAGCAAGACTCTGTCTCA 2900  
QY 2904 AAAAAAATAAAAAA 2917  
DB 2901 AAAAAAATAAAAAA 2914

## RESULT 2

US-10-104-047-997  
; Sequence 997, Application US/10104047  
; Patent No. 6943241  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. 6943241el full length cDNA  
; FILE REFERENCE: H1-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; CURRENT FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 997  
; LENGTH: 2759  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-104-047-997

Query Match 89.0%; Score 2598.2; DB 3; Length 2759;  
Best Local Similarity 95.2%; Pred. No. 0;  
Matches 2756; Conservative 0; Mismatches 3; Indels 136; Gaps 2;

QY 8 CTCTGCAGCCCTGAGCTGGGAAGAGAGCTACCTCGGAGGAGGCGGCGCGAGGCGGCG 67  
DB 1 CTCTGCAGCCCTGAGCTGGGAAGAGAGCTACCTCGGAGGAGGCGGCGGCGAGGCGGCG 60  
QY 68 GCGGATGAGAGGGGGGCGAGCGACCGCCCGGCTGGGAGGCGACCGCTAACCCCTGAC 127  
DB 61 GCGGATGAGAGGGGGGCGAGCGACCGCCCGGCTGGGAGGCGACCGCTAACCCCTGAC 120  
QY 128 CCCACCCACCCCTGACAAAGAGCTGGCGGGCGCTGGCCAGTCCCTGGGTGACCTT 187  
DB 121 CCACCCCA----- 128  
QY 188 CCTCGGATGAGAAATCCGCCCTCGGAGCATCTCTTCTCTCTAGGCTCTGAAGGCCCGG 247  
DB 129 -----CCCTCGGAGCATCTCTTCTCTCTAGGCTCTGAAGGCCCGG 170  
QY 248 GGAGCGTGAAGTCCAGTGCACCGCGGAGGAGGCTCGCTTTGTCAGTAAGGA 307  
DB 171 GGAGCGTGAAGTCCAGTGCACCGCGGAGGAGGCTCGCTTTGTCAGTAAGGA 230

QY 308 GGAGAGGCTGTCTCAGCTGCAGAGGGGTATCCCTGCTTCAAGCCAGTGCCTCTTCCAG 367  
DB 231 GGAGAGGCTGTCTCAGCTGCAGAGGGGTATCCCTGCTTCAAGCCAGTGCCTCTTCCAG 290  
QY 369 CTCCCATGAGGAGCCACCGAGCCACGCTCCGATGGAAGAAACCTGGAAGTGAAGGAGAT 427  
DB 291 CTCCCATGAGGAGCCACCGAGCCACGCTCCGATGGAAGAAACCTGGAAGTGAAGGAGAT 350  
QY 428 GGCAGGAGCAAGATCTTCCAGGCGCACTCCCAAGAGAGAGCGGGGTGGAACTGTCTGGCA 487  
DB 351 GGC----- 353  
QY 488 GGCCTGTGGAAGACACATCTCTCTCCCAACACGCTAAATTTCAACGAGCGCATCGTA 547  
DB 354 -----AAGACACATCTCTCTCCCAACACGCTAAATTTCAACGAGCGCATCGTA 404  
QY 548 AGAGGAGAGCGCTGTGGCCCCAGAGATCAACATTTCTCTGATCAGAGTGAAGGTTCC 607  
DB 405 AGAGGAGAGCGCTGTGGCCCCAGAGATCAACATTTCTCTGATCAGAGTGAAGGTTCC 464  
QY 608 TGCTCTCCGATGACTTCTTGGATACCCCTGATGACCTGGATATTAACGTGGATGACATCG 667  
DB 465 TGCTCTCCGATGACTTCTTGGATACCCCTGATGACCTGGATATTAACGTGGATGACATCG 524  
QY 668 AGACCCCGATGAGACCGACTCGCTGGAGTTCTCTGGGAAATGSCAACGAACTGGAGTGG 727  
DB 525 AGACCCCGATGAGACCGACTCGCTGGAGTTCTCTGGGAAATGSCAACGAACTGGAGTGG 584  
QY 728 AAGACGACACCCCTGTGGCCACCGCAAGAAATGSCCGGGGACAGCGGATCTATTG 787  
DB 585 AAGACGACACCCCTGTGGCCACCGCAAGAAATGSCCGGGGACAGCGGATCTATTG 644  
QY 788 GGGACGGACGACGAGGAGGACGCGCGCCCAACGCGGCTGTGGCGGACAGTGAATCA 847  
DB 645 GGGACGGACGACGAGGAGGACGCGCGCCCAACGCGGCTGTGGCGGACAGTGAATCA 704  
QY 848 TCGGGGAGCAAGAGCACCGTATAGACCTGCATGATCCGGCTTTACATGAAGTGGTCA 907  
DB 705 TCGGGGAGCAAGAGCACCGTATAGACCTGCATGATCCGGCTTTACATGAAGTGGTCA 764  
QY 908 CCCACGGAGGTTACTACGCGGAAGGCTCAACGCCATCATCGTCTTCCGAGCTGTCTCC 967  
DB 765 CCCACGGAGGTTACTACGCGGAAGGCTCAACGCCATCATCGTCTTCCGAGCTGTCTCC 824  
QY 968 TTCCAGACGACGACCTCCCGACTACCACTACATCATGAGAGAACCTCTTCTGTAGTCA 1027  
DB 825 TTCCAGACGACGACCTCCCGACTACCACTACATCATGAGAGAACCTCTTCTGTAGTCA 884  
QY 1028 TCAGCAGCTTAGAGCTCTCTGTGGCTGAGGACTACATGATCGTGTACCTGAACGGTGA 1087  
DB 885 TCAGCAGCTTAGAGCTCTCTGTGGCTGAGGACTACATGATCGTGTACCTGAACGGTGA 944  
QY 1088 CGCCCGGCGGAGGATGCTGGAATCGGCTGCTGAGAGGCTACCAAGTGTACCAAGATGCGAC 1147  
DB 945 CGCCCGGCGGAGGATGCTGGAATCGGCTGCTGAGAGGCTACCAAGTGTACCAAGATGCGAC 1004  
QY 1148 GAGGTTGCGGAAAAAAGCTGAAGTCTTGAATCATCGTCCACCCCTCGTGGTTCACTCGGA 1207  
DB 1005 GAGGTTGCGGAAAAAAGCTGAAGTCTTGAATCATCGTCCACCCCTCGTGGTTCACTCGGA 1064  
QY 1208 CTGTGCTGGCCATCTCTCGCCCTTTTCATCAGGCTCAAGTTCATCAACAGATCCAGTAG 1267  
DB 1065 CTGTGCTGGCCATCTCTCGCCCTTTTCATCAGGCTCAAGTTCATCAACAGATCCAGTAG 1124  
QY 1268 TGACAGCTTGAAGACCTGGAGCAACTCATCTCTATGGAACACGCTCCAGATCCAGACT 1327  
DB 1125 TGACAGCTTGAAGACCTGGAGCAACTCATCTCTATGGAACACGCTCCAGATCCAGACT 1184  
QY 1328 GCGTCTCGAATAAGCAAGAGGAAAGACTGAAGGCGGAGGAGGAGGCGGAGGCGGCGGCG 1387  
DB 1185 GCGTCTCGAATAAGCAAGAGGAAAGACTGAAGGCGGAGGAGGAGGCGGAGGCGGCGGCGG 1244



Db 1525 GTGGCTCCCTCTGCTCGACCCCTGGTGGATGAGCTGGATCAAGATTCAAGACC 1584  
Qy 1862 TTCTTTGCTTGTCAACCGCTCCAGGTTGGAGCCACAGACCCACCGCCCGGCTG 1921  
Db 1585 TTCTTTGCTTGTCAACCGCTCCAGGTTGGAGCCACAGACCCACCGCCCGGCTG 1644  
Qy 1922 GGTCTGGTCTTTCTTGTGCTTCCCTCAGAAATGCGGCTCAGACCTAGAGCTCAA 1981  
Db 1645 GGTCTGGTCTTTCTTGTGCTTCCCTCAGAAATGCGGCTCAGACCTAGAGCTCAA 1704  
Qy 1982 CCCCCCTATGAGGCGCAGCTCGGGTGTCTCTGAGCTCGGCTCAGACCTAGAGCTCAA 2041  
Db 1705 CCCCCCTATGAGGCGCAGCTCGGGTGTCTCTGAGCTCGGCTCAGACCTAGAGCTCAA 1764  
Qy 2042 CACACCCATGTTTTCAITTTGACCCCGCCCTTCTCGCTCAATAATGACACCCAGCTCT 2101  
Db 1765 CACACCCATGTTTTCAITTTGACCCCGCCCTTCTCGCTCAATAATGACACCCAGCTCT 1824  
Qy 2102 TTGAGAGGATCAGAGCCATTTGACAAGAGAGCGCTGCGCAACCAATCTTGTCTCGA 2161  
Db 1825 TTGAGAGGATCAGAGCCATTTGACAAGAGAGCGCTGCGCAACCAATCTTGTCTCGA 1884  
Qy 2162 TTGCAAAATGACACCCAGTAATCTAGAACATTTCTCAGCCCTTTAACTCAGATGTCAA 2221  
Db 1885 TTGCAAAATGACACCCAGTAATCTAGAACATTTCTCAGCCCTTTAACTCAGATGTCAA 1944  
Qy 2222 GCCACCGGCGCAACCCCGTCAATACCTCCACCAAGGAATGAGATGTGGACCTCAGTG 2281  
Db 1945 GCCACCGGCGCAACCCCGTCAATACCTCCACCAAGGAATGAGATGTGGACCTCAGTG 2004  
Qy 2282 CTCCTCCCAACCCAGCTCAGGCTGGGACACGCGCAACGCTGTTCCGGTTGGAAACAGCAGA 2341  
Db 2005 CTCCTCCCAACCCAGCTCAGGCTGGGACACGCGCAACGCTGTTCCGGTTGGAAACAGCAGA 2064  
Qy 2342 GGTCTAGAACTGGCTCTGAATAGCAGACCTAGCAGAGGAATGATCAGGGTATCGGG 2401  
Db 2065 GGTCTAGAACTGGCTCTGAATAGCAGACCTAGCAGAGGAATGATCAGGGTATCGGG 2124  
Qy 2402 CGTTTCAGTGTTCAGAACTGTTTCGGGAAGATTAATTCAGTGGCTGGCCGCGCACCC 2461  
Db 2125 CGTTTCAGTGTTCAGAACTGTTTCGGGAAGATTAATTCAGTGGCTGGCCGCGCACCC 2184  
Qy 2462 TGCAATTCAGGCTGGACACGCGGTTCTTGTTCGGGAGGCAAAATTCCTTAGGAAAAG 2521  
Db 2185 TGCAATTCAGGCTGGACACGCGGTTCTTGTTCGGGAGGCAAAATTCCTTAGGAAAAG 2244  
Qy 2522 AAGA CAGACTTTCTAATGTGTCAAATGCGGATCACTGGTCAAGTGAATCTAGAAC 2581  
Db 2245 AAGA CAGACTTTCTAATGTGTCAAATGCGGATCACTGGTCAAGTGAATCTAGAAC 2304  
Qy 2582 ACTGAGCTCCCTGCTCTCGAAGTATTTAAGAAAGGCTGGCCAGGACGATGCTCAC 2641  
Db 2305 ACTGAGCTCCCTGCTCTCGAAGTATTTAAGAAAGGCTGGCCAGGACGATGCTCAC 2364  
Qy 2642 GCCTGTAAATCCAGACTTTGGGAGGCGGAGGCGGATCACTGAGGTGAGGAGTTTG 2701  
Db 2365 GCCTGTAAATCCAGACTTTGGGAGGCGGAGGCGGATCACTGAGGTGAGGAGTTTG 2424  
Qy 2702 AAGA CAGCTTGGCCAAATGTTGAAACCTATCTCTAATAAATAAATAAATAAATAAATAA 2761  
Db 2425 AAGA CAGCTTGGCCAAATGTTGAAACCTATCTCTAATAAATAAATAAATAAATAAATAA 2484  
Qy 2762 GGTGTGTGCGAGTGTCTGTAATCCAGCTACTTGGAGGCTGAGGATGAGATCACTT 2821  
Db 2485 GGTGTGTGCGAGTGTCTGTAATCCAGCTACTTGGAGGCTGAGGATGAGATCACTT 2544  
Qy 2822 AAACCTGAGAGGCGAGGTTTACAGTGAAGCAAGATCGTCCACTGATTTCCAGCTGGGC 2881  
Db 2545 AAACCTGAGAGGCGAGGTTTACAGTGAAGCAAGATCGTCCACTGATTTCCAGCTGGGC 2604  
Qy 2882 GACAGGCAAGACTCTGTCTC 2902

Db 2605 GACAGAGCAAGACTCTGTCTC 2625

## RESULT 4

US-09-174-937-3  
; Sequence 3, Application US/09174937  
; Patent No. 6756212  
; GENERAL INFORMATION:  
; APPLICANT: Khodadoust, Mehran  
; TITLE OF INVENTION: Isolated Proteins and Nucleic Acid Molecules Having  
; TITLE OF INVENTION: Homology to NIP2 Proteins and Uses Thereof  
; FILE REFERENCE: mn1061  
; CURRENT APPLICATION NUMBER: US/09/174,937  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1113  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1113)  
US-09-174-937-3

Query Match 38.1%; Score 1113; DB 3; Length 1113;

Best Local Similarity 100.0%; Pred. No. 5.8e-262; Mismatches 0; Indels 0; Gaps 0;  
Matches 1113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 373 ATGGGAGCACCGAAGCCACGCTCCGGATGGAAAACGTGGAGCTGAAGAGGAATGGCAG 432

Db 1 ATGGGGACACCGAAGCCACGCTCCGGATGGAAAACGTGGAGCTGAAGAGGAATGGCAG 60

Qy 433 GACGAAGATCTTCCAGAGCCACTCCAGAGAGAGCGGGGTGGAATCTGCTTGGCAGCCCG 492

Db 61 GACGAAGATCTTCCAGAGCCACTCCAGAGAGAGCGGGGTGGAATCTGCTTGGCAGCCCG 120

Qy 493 GTGGAAGACACATCTCTCCCAACAGCTTAATTTCAACGGAGCGCATCGTAAAGAGG 552

Db 121 GTGGAAGACACATCTCTCCCAACAGCTTAATTTCAACGGAGCGCATCGTAAAGAGG 180

Qy 553 AAGACGCTGGTGGCCCGCAGAGATCAACATTTCTCTGGATCAGAGTGAAGGGTCCCTGCTG 612

Db 181 AAGACGCTGGTGGCCCGCAGAGATCAACATTTCTCTGGATCAGAGTGAAGGGTCCCTGCTG 240

Qy 613 TCCGATGATCTTCTGGATACCCCTGATGACCTGGATATTAACGTGGATGACATCGAGACC 672

Db 241 TCCGATGATCTTCTGGATACCCCTGATGACCTGGATATTAACGTGGATGACATCGAGACC 300

Qy 673 CCCGATGAGACCGACTCCCTGGAGTTCCCTGGGATGGCAACGAACTGGAGTGGGAAGAC 732

Db 301 CCCGATGAGACCGACTCCCTGGAGTTCCCTGGGATGGCAACGAACTGGAGTGGGAAGAC 360

Qy 733 GACACCCCGTGGCCACCGCCAAAGCAATGCGCGGGAGCAGCGCGGATCTATTTCGGGAC 792

Db 361 GACACCCCGTGGCCACCGCCAAAGCAATGCGCGGGAGCAGCGCGGATCTATTTCGGGAC 420

Qy 793 GGCACGACGAGAGCGGAGCGCGCCAAACGGGGCCCTGTGGCGGACAGTGAATCATCGGG 852

Db 421 GGCACGACGAGAGCGGAGCGCGCCAAACGGGGCCCTGTGGCGGACAGTGAATCATCGGG 480

Qy 853 GAGCAAGAGCACCGTATAGACCTGCACATGATCCGGCCCTTACNTGAAGTGGTCAACCCAC 912

Db 481 GAGCAAGAGCACCGTATAGACCTGCACATGATCCGGCCCTTACNTGAAGTGGTCAACCCAC 540

Qy 913 GGAGGGTACTACGGCGAAGGCTCAACCGCATCATCGTCTTCGACAGCTGCTTCCTTCCA 972

Db 541 GGAGGGTACTACGGCGAAGGCTCAACCGCATCATCGTCTTCGACAGCTGCTTCCTTCCA 600

Qy 973 GACAGAGCTCCCGGACTACCACTACATCATATGAGGAACTCTTCTCTAGCTCATCAGC 1032

Db 601 GACAGAGCTCCCGGACTACCACTACATCATATGAGGAACTCTTCTCTAGCTCATCAGC 660

QY 1033 AGCTTAGAGTCTCTGGTGGCTGAGGACTACATGATCGTGTACTTGAACGGTGCACGCCC 1092  
Db 661 AGCTTAGAGTCTCTGGTGGCTGAGGACTACATGATCGTGTACTTGAACGGTGCACGCCC 720  
QY 1093 CGCGGAGGATGCTGGAATCGGCTGGCTGAAGAAGTGTACAGATGATCGACCGGAGG 1152  
Db 721 CGCGGAGGATGCTGGAATCGGCTGGCTGAAGAAGTGTACAGATGATCGACCGGAGG 780  
QY 1153 TTGCGGAAACCTGAAGTCTTGATCATGCTGCACCCCTCGTGGTTTCATTGCGACTGTG 1212  
Db 781 TTGCGGAAACCTGAAGTCTTGATCATGCTGCACCCCTCGTGGTTTCATTGCGACTGTG 840  
QY 1213 CTGGCCATCTCTCGCCCTTTCATCAGCGTCAAGTTTCATCAACAAGATCCAGTACGTGCAC 1272  
Db 841 CTGGCCATCTCTCGCCCTTTCATCAGCGTCAAGTTTCATCAACAAGATCCAGTACGTGCAC 900  
QY 1273 AGCTTGGAGACCTGGAGCACTCATCCCTATGGAACAGTCCAGATCCAGACTCGTC 1332  
Db 901 AGCTTGGAGACCTGGAGCACTCATCCCTATGGAACAGTCCAGATCCAGACTCGTC 960  
QY 1333 CTGCAATACGAAGAGAACTGAAGGCGGAGGAGGAGGCGGAGGCGGAGGCGGAGG 1392  
Db 961 CTGCAATACGAAGAGAACTGAAGGCGGAGGAGGAGGCGGAGGCGGAGGCGGAGG 1020  
QY 1393 TTTGTGTCGCCAGTCTGAAGAGAGCCAGAGTGGCCACCAAGTGGAAAACAGGTCTGCT 1452  
Db 1021 TTTGTGTCGCCAGTCTGAAGAGAGCCAGAGTGGCCACCAAGTGGAAAACAGGTCTGCT 1080  
QY 1453 CTGCTCTCAGAGATCAGGAAACAGCATGTCC 1485  
Db 1081 CTGCTCTCAGAGATCAGGAAACAGCATGTCC 1113

## RESULT 5

US-09-621-976-14512  
; Sequence 14512, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 14512  
; LENGTH: 322  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-14512

Query Match 10.6%; Score 309.8; DB 3; Length 322;  
Best Local Similarity 98.8%; Pred. No. 4.8e-66;  
Matches 319; Conservative 3; Mismatches 0; Indels 1; Gaps 1;  
QY 2575 TAGAAGCACTGAGCTCCCTGCTCTCTGGAAGTATTTAAGAAAGGCTGGCCAGGCACGAT 2634  
Db 1 TAGAAGCACTGAGCTCCCTGCTCTCTGGAAGTATTTAAGAAAGGCTGGCCAGGCACGAT 60  
QY 2635 GGCTACGCGCTGTAATCCAGACTTTTGGGAGGCGGAGGAGGCGGATCACTGAGGTGAG 2694  
Db 61 GGCTACGCGCTGTAATCCAGASTTTTGGGAGGCGGAGGAGGCGGATCACTGAGGTGAG 120  
QY 2695 GAGTTTGAACAAGCTCGCCAAACATGGTGAAACCTCATCTCTATAAAAAATACAAAAAT 2754  
Db 121 GAGTTTGAACAAGCTCGCCAAACATGGTGAAACCTCATCTCTATAAAAAATACAAAAAT 180  
QY 2755 TAGCCAGGCGTGGTGGCAGTGTCTTAATCCAGCTACTTGGAGGCTGAGGCATGAGA 2814  
Db 181 TAGCCAGGCGTGGTGGCAGTGTCTTAATCCAGCTACTT-GGAGGCTGAGGCATGAGA 239

QY 2815 ATCACTTAAACCTGAGAGCGAGAGGTTTACAGTAGCCCAAGATCGTGCCACTGCAATTCAG 2874  
Db 240 ATCACTTAAACCTGAGAGCGAGAGGTTTACAGTAGCCCAAGATCGTGCCACTGCAATTCAG 299  
QY 2875 COTGGCGACAGAGCAAGACTCT 2897  
Db 300 COTGGCGACAGAGCAAGACTCT 322

## RESULT 6

US-09-174-937-6  
; Sequence 6, Application US/09174937  
; Patent No. 6756212  
; GENERAL INFORMATION:  
; APPLICANT: Khodadoust, Mehran  
; TITLE OF INVENTION: Isolated Proteins and Nucleic Acid Molecules Having  
; TITLE OF INVENTION: Homology to NIP2 Proteins and Uses Thereof  
; FILE REFERENCE: m11061  
; CURRENT APPLICATION NUMBER: US/09/174,937  
; CURRENT FILING DATE: 1998-10-19  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 966  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(966)  
US-09-174-937-6

Query Match 10.4%; Score 302.2; DB 3; Length 966;  
Best Local Similarity 62.5%; Pred. No. 5.2e-64;  
Matches 547; Conservative 0; Mismatches 28; Indels 45; Gaps 3;  
QY 469 GGGGTGGAACTGCTTTGGCAGCCCGGTGGAAGACACATCTCTCTCCCAACACGTAAT 528  
Db 97 GAGGGCGTGTGAGTCCCGAGTGTGACAGATGAGGCGCTGAACCTCTTAATTTCTGGAT 156  
QY 529 TTCAACGGAGCGCATCGTAAGAGGAGAGCGTGGTGGGCCCCAGAGATCAACATTTCTCTG 588  
Db 157 CTTAATGACATCTCATCTCTCGGAGATCAAGCTCAAGCCCCCAATATCAATCTTTCTCTG 216  
QY 589 GATCAGATGAGGGTCCCTGCTGCCATGACTTCTTGGATACCCCTGATGACCTGGAT 648  
Db 217 GACCAAGTGAAGATCTATTCTCTCATGATTAATTTGGACAGTCCAGATGAATTCAC 276  
QY 649 ATTAACTGGATGACATCGAGACCCCGATGAGACCGACTCGCTGGAGTTCTCTGGGAAT 708  
Db 277 ATCAATGTGGATGAACCTTGTATACCCCGATGAAGCAGATTTCTTTTGTAGTACACTGG 332  
QY 709 GGCAAGCACTGGAGTGGGAGACGACACCCCGTGGCCACCGCCCAAGAACATGCCCGGG 768  
Db 333 -----CCATGATCCCAAGCAAGAAATTCGGC 363  
QY 769 GACAGCGCGATCTATTGGGAGCGGACGACGAGGAGGAGCGGAGCGCCCAACGGGCGC 828  
Db 364 CAAGAGTCAGAGTCTATTCCAGATATATAGCCGAGAGAGGAGGAGGAGCAAC---CG 420  
QY 829 CTGTGGCGGACAGTATCATCGGGAGCAAGAGCACCGTATAGACCTTGCAATGATCCGG 888  
Db 421 CTTTGGATGACAGTGGTCAATTGGAGAACAGAGCAGCGCATTTGACATGAAGTTCATCGAG 480  
QY 889 CCTTACATGAAGTGGTCAACCCACGG-----AGGGTACTACGGCGAGGCGCTCAAC 939  
Db 481 CCCTACAGGAGTCAATTTCTCACGGAGGAGATTCCAGGATCTATGGGGACGGTCTAAAT 540  
QY 940 GCATCATCTGCTTCGCGAGCTGCTTCTTCCAGACAGCAGCTCCCGACTTACCACATAC 999  
Db 541 GCCATCATGTTGTTGGCGCTGTTTCTGCCAGACAGAGTCCGGCGGATTAACCATAT 600  
QY 1000 ATCATGGAGAACCTTCTCTGTATCGTATCATCAGCAGCTTAGAGCTCTCTGGTGGCAGGAC 1059



Db	601	GT	CATGG	AAATCTTTTC	TATATGTA	TAAGTACTT	TAGAGTTG	TGAGTGTG	AGAGAC	660
Qy	1060	TAC	TGATG	TGTTACT	TGAACGG	TGCCAC	CGCCCCG	CGGAGG	TGCTTGAATCGGCTGG	1119
Db	661	TAT	TGATGTTG	TGTTACTT	TGAATG	TGTGCAAC	CCCAAG	AGGAGG	TGCCAGGCTAGGCTGG	720
Qy	1120	CTG	AAGAAGT	CTAC	CAGATG	ATCGAC	CGGAGG	TTCG	GGAACCTGGAAGTCTCTTGATC	1179
Db	721	ATG	AGAAATG	CTAC	CAGATG	ATTCAC	AGCGTTG	AGGAGAA	TTTGAATCATTTCATC	780
Qy	1180	ATC	GTCAC	CCCTCGT	GGTTCA	TTTGG	ACTGTG	TGCGCC	ATCTCTGCCCTTTTCATCAGC	1239
Db	781	ATT	GTTTC	ATCC	ATCTTGG	TTTCAT	CAGAA	CAATCTCTG	CTGTGCACGACCTTTTAAAGT	840
Qy	1240	GT	CAAGTTT	CATCA	CAAGAT	CCAGT	AGTGC	CACAGCTT	TGGAAGACCTTGGAGCAACTCATC	1299
Db	841	TC	AAAATTC	AC	GAGTAA	NTAATAT	TGTC	ATAGCTTAT	CAGAACTCAGTGGGCTGATC	900
Qy	1300	CCT	ATGGA	CAAC	GTGCC	AGATCCC	CAGACTGG	GTCT	1334	
Db	901	CCA	TGATG	TTCAT	CTCC	ACATTC	CA	GAGAGAT	CAT	935

## RESULT 7

```

US-09-174-937--4
; Sequence 4, Application US/09174937
; Patent No. 6758212
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: Isolated Proteins and Nucleic Acid Molecules Having
; TITLE OF INVENTION: Homology to NIP2 Proteins and Uses Thereof
; FILE REFERENCE: mm1061
; CURRENT APPLICATION NUMBER: US/09/174,937
; CURRENT FILING DATE: 1998-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 4235
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (23)..(988)
US-09-174-937-4

```

## RESULT 8

```

US-09-949-016-12513/C
; Sequence 12513, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12513
; LENGTH: 103987
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12513

```

Qy	829	CTGTGGCGGACAGTGTATCATCGGGGACGAAGAGCACCGGTATAGACCTGCACTGATGCCG 888
Db	443	CTTTGGATGACAGTGTCTATTGGAGAAACAAGAGCAGCGCATTTGACATGAAGGTTCATCGAG 502
Qy	889	CCTTTACATGAAAGTGTCTACCCACGG-----AGGTTACTTACGGCGAAGCGCTCAAC 939
Db	503	CCCTACAGGAGAGTCTATTTCTCACGGAGGAGATTTCAGGATACTATATGGGACGCGTCTTAAT 562
Qy	940	GCCATCATCGTCTTTCCGACGCTGCTTCTTCCAGACAGCAGAGCCTCCCGACTTACCACTAC 999
Db	563	GCCATCATTTGTGTTGCCGCGCTGTTTCTGCCAGACAGCAGTGGGGCGGATTACCACTAT 622
Qy	1000	ATCATGGAGAACCTCTTCTGTAGCTCATCAGCAGCTTAGAGCTCTCGTGGCTGAGGAC 1059
Db	623	GTCAATGGAAATCTTTTCTTATATGTAATAAGTACTTTTAGAGTTTGATGTTAGCTGAAGAC 682
Qy	1060	TACATGATCGTGTACCTGAAACGGTGTGCACGCCCGCGGAGGATGCTCGGAATCGCGTGG 1119
Db	683	TATATGATTTGTGTACTTTGAATGGTGTCAAACCCCAAGAAGGAGGATGCGAGGGCTAGGCTGG 742
Qy	1120	CTCAAGAAGTGCTTACCAGATGATCGACCGGAGGTTTCGGGAAAAAAGTGAAGTCTTGATC 1179
Db	743	ATGAAGAAATGCTTACCAGATGATTTGACAGACGCGTTTGAGGAAGAAATTTGAATTCATTTCATC 802
Qy	1180	ATCGTCCACCCCTCGTGGTTTCATTCGGGACTGTGCTGGCCATCTCTCGCCCTTTTCATCAGC 1239
Db	803	ATTGTTTCATGCATCTTGGTTTCATCAGAAACATCTTGTCTGTGACACGACTTTTATAAGT 862
Qy	1240	GTCAAGTTTCATCAACAAGATCCAGTACGTCGACAGCTTTGGAAAGACCTTGGAGCAATCTCATC 1299
Db	863	TCAAAATTCAGCAGTAAATTTAAATATGTCAATAGCTTATCAGAACTCTAGTGGGCTGATC 922
Qy	1300	CCTATGGACACGTCACAGATCCAGATCCGAGACTGGGTCCT 1334
Db	923	CCAAATGGATTTGATCCACATTTCCAGAGACATCAT 957

	Query Match	8.3%	Score 241.8;	DB 3;	Length 103987;
	Best Local Similarity	89.2%;	Pred. No. 1.8e-48;		
	Matches 272;	Conservative 0;	Mismatches 32;	Indels 1;	Gaps 1;
Qy	2614	AAAGGCTGGGCAGGCA	CGATGGCTCACGCTGTAA	TCCAG-AC	TTTTGGGAGGCGGAGG 2672
Db	47908	AAAGTCTGGGCTGGGCG	GGTGGCTCATGCTGTAA	TCCAGCAGCT	TTTGGTGGCTGAGG 47849
Qy	2673	CAGGCGGATCA	CCCTGAGGTGACGAGTT	TGAGAACGCTGG	GCACATGGTGAACCTTCA 2732
Db	47848	CAGGCAGATCT	CTCTGAGGTTCAGAGTT	TGAGACTAGCTGG	GCACATGGTGAACCTTCA 47789



NAME/KEY: misc feature  
LOCATION: (1)...(193169)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15091

Query Match 8.3%; Score 241.2; DB 3; Length 193169;  
Best Local Similarity 85.2%; Pred. No. 3.2e-48;  
Matches 281; Conservative 0; Mismatches 48; Indels 1; Gaps 1;  
2590 CCTGTCTCTGGAAGTATTAAAGAAAGGCTGGGCCAGGCAACATGGCTCAGCCCTGTAA 2649  
Db |||||  
76929 CCCATCTCTAAATATATAATTAATAAACAAGCCAGGTATGCTGCTCAGACCTGTAA 76870  
Qy |||||  
2650 TCCAG-ACCTTGGAGGCGGAGCGAGCGGAGTCACTGAGGTGAGGAGTTTGAGAACAG 2708  
Db |||||  
76869 TCCAGCACCTTGGAGGCGGAGCGAGGTGAGTCACTGAGGCGAGGATGAGACCAAG 76810  
Qy |||||  
2709 CTGGGCCAACATGTAACCTCATCTCTACTATAAATAAATAAATAAATAAATAAATAA 2768  
Db |||||  
76809 CTGAACAACATGTAACCTCATCTCTACTATAAATAAATAAATAAATAAATAAATAA 27650  
Qy |||||  
2769 GCGAGGTGCTGTAATCCCACTACTTGGAGGCTGAGGATGAGGATCACTTAAACCTG 2828  
Db |||||  
76749 GCGGGTGTCTGCAATCCCGCTACTCGGAGGCTGAGGATGAGGATCACTTGAACCCG 76690  
Qy |||||  
2829 AGAGGAGAGGTTACAGTGAAGCCAGATCGTGCCACTGCAATCCAGCTGGGCGACAGAG 2888  
Db |||||  
76689 GAGGCGAGAGTTGCAAGTGAAGGAGTGTGCTACTGCAATCCAGGCTGGGCGACAGAG 76630  
Qy |||||  
2889 CAAGACTGTCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2918  
Db |||||  
76629 TGAGACTTTGCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 76600

## RESULT 12

US-09-949-016-17521  
Sequence 17521, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 17521  
LENGTH: 87734  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-17521

Query Match 8.2%; Score 240.4; DB 3; Length 87734;  
Best Local Similarity 88.1%; Pred. No. 3.7e-48;  
Matches 273; Conservative 0; Mismatches 36; Indels 1; Gaps 1;  
2610 AAGAAAGGCTGGCCAGGCAAGATGCTCAGCCCTGTAATCCAG-ACCTTGGAGGCG 2668  
Db |||||  
70886 AAGAAAGGCTGGCTGGTGGCGGCTGCTGCGCCCTGTAATCCAGCACTTTGGAGTCC 70945  
Qy |||||  
2669 GAGGAGGCGGATCACTGAGGTGAGGAGTTTGAGAACAGCTGGCCACATGTTGAAC 2728  
Db |||||  
70946 GAGGAGGCGGATCACTGAGGTGAGGAGTTTGAGAACAGCTGGCCACATGTTGAAC 71005  
Qy |||||  
2729 CTCATCTCTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2788  
Db |||||

Db 71006 CTCGTCTCTACTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 71065  
Qy 2789 GCTACTTGGAGGCTGAGGATGAGATCACTTAAACCTGAGAGGCGAGGTTACAGTGA 2848  
Db |||||  
71066 GCTACTTGGAGGCTGAGGATGAGATCACTTAAACCTGAGAGGCGAGGTTACAGTGA 71125  
Qy 2849 GCCAAGATCGTCCCACTGCAATCCAGCCCTGGGCGGACAGAGCAAGACTCTGTCTCAAAAAA 2908  
Db |||||  
71126 GCGAGATGGCGCACTGCACTCCAGCCCTGGGCGGACAGAGCAAGACTCTCAAAAAA 71185  
Qy 2909 AAAAAAATAA 2918  
Db 71186 AAAAAAATAA 71195

## RESULT 13

US-09-949-016-13408/c  
Sequence 13408, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13408  
LENGTH: 49378  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-13408

Query Match 8.2%; Score 239.8; DB 3; Length 49378;  
Best Local Similarity 75.5%; Pred. No. 4.2e-48;  
Matches 324; Conservative 0; Mismatches 102; Indels 3; Gaps 2;  
2491 TGTTGGGAGGCAATTTCCCTAGAGAAAGAGACACACTTTCTTAATGTGCTCAAT 2550  
Db |||||  
4821 TGACTGTCTTAATGTTGCTGAGAGAAACACAAACAAACAAACAAACAAACAAACAA 4762  
Qy 2551 GCGGATCACTGTCAGATGGACTCTAGAGCACTGAGCTCCCTGCTCTCGGAAGTATTA 2610  
Db |||||  
4761 TAAGCAATGATCAATAAATACTAAGAAAGACAGTGT--TATACCTTAAACATCTTTA 4704  
Qy 2611 AAGAAAGGCTGGCCAGGCAAGATGCTCAGCCCTGTAATCCAG-ACCTTGGAGGCGG 2669  
Db |||||  
4703 AAGAGGTTGAGGCGGCGGCGGCTGCTCAGCTGTAATCCAGCACTTTGGAGGCGG 4644  
Qy 2670 AGGAGGCGGATCACTGAGGTGAGGAGTTTGAGAACAGCTGGCCACATGTTGAAACC 2729  
Db |||||  
4643 AGGAGGCGGATCACTGAGGTGAGGAGTTTGAGAACAGCTGGCCACATGTTGAAACC 4584  
Qy 2730 TCATCTCTACTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2789  
Db |||||  
4583 CCGTCTCTATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4524  
Qy 2790 CTACTTGGAGGCTGAGGATGAGATCACTTAAACCTGAGAGGCGAGAGGTTTACAGTGA 2849  
Db |||||  
4523 CTACTGAGAGGCTGAGTCAAGAAATCACTTGAACCTGGAGAGAGAGGTTGAGTGA 4464  
Qy 2850 CCAAGATCGTCCCACTGCAATTCAGCCCTGGGCGGACAGAGCAAGACTCTGTCTCAAAAAA 2909  
Db |||||  
4463 CCAAGATGTCTACTGCTCACTCCAGCCCTGGGCGGACAGAGCAAGACTCTCAAAAAA 4404  
Qy 2910 AAAAAAATAA 2918

Db 4403 AAAAGAAA 4395

## RESULT 14

US-09-949-016-17290  
; Sequence 17290, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17290  
; LENGTH: 11929  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17290

Query Match 8.2%; Score 239.6; DB 3; Length 11929;

Best Local Similarity 87.3%; Pred. No. 2.7e-48;

Matches 274; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 2605 TATTTAAGAAAGCGCTGGCGCAGCATGGCTCAGCGCTGTAATCCAG-ACCTTGGG 2663

Db 6427 TAAAAAATAAATAATGGCCGGCGAGTGGCTCACCTGTAATCCAGCACTTTGG 6486

QY 2664 AGCCGAGGCGCGGATCACTTGAGGTGAGGAGTTTGAGAAACAGCCTGGCCAAATGGT 2723

Db 6487 AGCCGAGGCGCGGATCACTTGAAGTCAGGAGTTTCGAGACCAGCCTGGCCAAATGGT 6546

QY 2724 GAAACCTCATCTTCTACTAAAAATACAAAAATTAGCCAGCGGTGGTGGCAGGTGCTGTAA 2783

Db 6547 GAAACCTCATCTTCTACTAAAAATACAAAAATTAGCCAGCGGTGGTGGCAGGTGCTGTAA 6606

QY 2784 TCCAGCTACTTGGGAGGCTGAGCATGAGAACTTAAACCTGAGAGCGAGGTGAC 2843

Db 6607 TCCAGCTACTTGGGAGGCTGAGCATGAGAACTTAAACCTGAGAGCGAGGTGAC 6666

QY 2844 AGTAGCCCAAGATCGTGCCACTGCAATTCAGCCTGGGCGACAGAGCAAGACTCTGTCTCA 2903

Db 6667 AGTAGCCCAAGATCGTGCCACTGCAATTCAGCCTGGGCGACAGAGCAAGACTCTGTCTCA 6726

QY 2904 AAAAAAAAAAAAAA 2917

Db 6727 AAAAAAAAAAAAAA 6740

## RESULT 15

US-09-949-016-14964/c  
; Sequence 14964, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14964  
; LENGTH: 24221  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(24221)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14964

Query Match 8.2%; Score 239.4; DB 3; Length 24221;

Best Local Similarity 88.0%; Pred. No. 4e-48;

Matches 272; Conservative 0; Mismatches 36; Indels 1; Gaps 1;

QY 2610 AAGAAAGCGCTGGCGCAGCATGGCTCAGCGCTGTAATCCAG-ACCTTGGGAGGCC 2668

Db 23246 AAGCGTAGGGGTGGCGAGCATGATGGCTCAGCGCTGTAATCCGCGCACTTTGGGAGGCC 23187

QY 2669 GAGGAGGCGGATCACCTGAGGTGAGGAGTTTGAGAAACAGCCTGGCCAAACATGGTGAAC 2728

Db 23186 AAGGAGGTTGGATCACCTGAGGTGAGGAGTTTCGAGACCAGCCTGGCCAAACATAGTGAAC 23127

QY 2729 CTCATCTCTACTAAAAATACAAAAATTAGCCAGCGGTGGTGGCAGGTGCTGTATCCCA 2788

Db 23126 CCGCTCTCTCTAAAAATACAAAAATTAGCCAGGCGATGGTGGCAGCTGCCTGTATCTTT 23067

QY 2789 GCTACTTTGGGAGGCTGAGGCGATGAGAACTTAAACCTGAGAGCGAGGTTTACAGTGA 2848

Db 23066 GCTACTTTGGGAGGCTGAGGCGAGGAGATCACTGAACTCCGCGAGGCGGAGGTTTACAGTGA 23007

QY 2849 GCCAAGATCGTGCCCACTGCAATTCAGCCTGGGCGGACAGAGCAAGACTCTGTCTCAAAAA 2908

Db 23006 GCCAAAAATCGCAACCACTGCACTCCAGCCTGGGCGGACAGAGCAAGACTCTCATCTCAAAAA 22947

QY 2909 AAAAAAAAAA 2917

Db 22946 AAAAAAGAA 22938

Search completed: March 5, 2006, 04:34:12

Job time : 506 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 03:44:45 ; Search time 2058 Seconds  
(without alignments)  
11724.998 Million cell updates/sec

Title: US-10-699-941-3

Perfect score: 2918

Sequence: 1 gccgagcctctgcagccct.....tctcaaaaaaaaaaaaaa 2918

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:\*

- 1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 3: /cgn2\_6/prodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 4: /cgn2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 5: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 6: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 8: /cgn2\_6/prodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 10: /cgn2\_6/prodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2918	100.0	2918	7	US-10-699-941-3
2	2916.4	99.9	2918	7	US-10-699-941-3
3	2598.2	89.0	2759	6	US-10-104-047-997
4	1438.2	49.3	47219	7	US-10-699-941-11
5	1438.2	49.3	59884	7	US-10-699-941-11
6	1325.8	45.4	2625	6	US-10-104-047-986
7	854.6	29.3	2166	7	US-10-699-941-1
8	675	23.1	703	9	US-10-450-763-16127
9	420.4	14.4	485	3	US-09-918-995-2340
10	330.4	11.3	5656	8	US-10-723-860-4072
11	330.4	11.3	5656	9	US-10-756-149-4047
12	327.6	11.2	11690	3	US-10-487-593-2
13	318.2	10.9	5654	3	US-09-873-367C-644
14	318.2	10.9	5654	3	US-09-968-007A-349
15	318.2	10.9	5654	3	US-09-968-007A-729
16	318.2	10.9	5654	9	US-10-843-641A-644
17	318.2	10.9	5654	9	US-10-843-641A-6819
18	318.2	10.9	5654	9	US-10-843-641A-7199
19	315.8	10.8	2347	6	US-10-220-891-19
20	302.2	10.4	4378	8	US-10-357-930-22138
21	302.2	10.4	4378	8	US-10-357-930-28000
22	267	9.2	786	5	US-10-027-632-152869
23	267	9.2	786	6	US-10-027-632-152869

Sequence 515679, A  
Sequence 13015, A  
Sequence 34, Appl  
Sequence 5729, Ap  
Sequence 594866, A  
Sequence 263139, A  
Sequence 263139, A  
Sequence 108, App  
Sequence 594865, A  
Sequence 55092, A  
Sequence 8024, Ap  
Sequence 482665, A  
Sequence 563212, A  
Sequence 251984, A  
Sequence 251984, A  
Sequence 5691, Ap  
Sequence 17762, A  
Sequence 621024, A  
Sequence 84578, A  
Sequence 456, App  
Sequence 664, App  
Sequence 723891, A

24 243.6 8.3 623 4 US-09-925-065A-515679  
25 243.6 8.3 660 3 US-09-814-353-13015  
26 241.8 8.3 96595 6 US-10-034-650-34  
27 241.6 8.3 403035 7 US-10-741-601-5729  
28 241.4 8.3 560 4 US-09-925-065A-594866  
29 240.6 8.2 548 5 US-10-027-632-263139  
30 240.6 8.2 548 6 US-10-027-632-263139  
31 240 8.2 100608 8 US-10-417-375-108  
32 239.8 8.2 560 4 US-09-925-065A-594865  
33 239.6 8.2 628 4 US-09-925-065A-55092  
34 239.6 8.2 32249 3 US-09-764-891-8024  
35 239.2 8.2 562 4 US-09-925-065A-482665  
36 239.2 8.2 611 4 US-09-925-065A-563212  
37 239.2 8.2 1046 5 US-10-027-632-251984  
38 239.2 8.2 1046 5 US-10-027-632-251984  
39 239.2 8.2 67126 7 US-10-741-601-5691  
40 239.2 8.2 67126 8 US-10-741-601-17762  
41 239 8.2 483 4 US-09-925-065A-621024  
42 239 8.2 1082 4 US-09-925-065A-84578  
43 239 8.2 301477 7 US-10-322-281-456  
44 238.8 8.2 59856 5 US-10-087-192-664  
45 238.6 8.2 2649 4 US-09-925-065A-723891

## ALIGNMENTS

## RESULT 1

US-10-699-941-3  
; Sequence 3, Application US/10699941  
; Publication No. US20040146900A1  
; GENERAL INFORMATION:  
; APPLICANT: Burmeister, Margit  
; TITLE OF INVENTION: Ataxia Associated Gene and Protein  
; FILE REFERENCE: UM-08441  
; CURRENT APPLICATION NUMBER: US/10/699,941  
; CURRENT FILING DATE: 2003-11-03  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 2918  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-699-941-3

Query Match 100.0%; Score 2918; DB 7; Length 2918;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2918; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCGAGCCTCTGCCAGCCCTGAGCTGGGAAGAGAGTACTCGGAGGCGGGCGCGCA 60  
Db 1 GCCGAGCCTCTGCCAGCCCTGAGCTGGGAAGAGAGTACTCGGAGGCGGGCGCGCA 60  
Qy 61 GCGCGGCGCGGATGAGAGGGGGCGCAGCCGAGCCCGCGCTGGGAGCCACCGCTAAC 120  
Db 61 GCGCGGCGCGGATGAGAGGGGGCGCAGCCGAGCCCGCGCTGGGAGCCACCGCTAAC 120  
Qy 121 CCTGCAACCCCAACCCCTCTGCAAAAAGAGCTGGCGGGCGCTGGCCACGTCGCTGGG 180  
Db 121 CCTGCAACCCCAACCCCTCTGCAAAAAGAGCTGGCGGGCGCTGGCCACGTCGCTGGG 180  
Qy 181 TGACCTTCTCTGGATGAGATCGCCCTCTGCAAAAAGAGCTGGCGGGCGCTGGCCACGTCGCTGGG 240  
Db 181 TGACCTTCTCTGGATGAGATCGCCCTCTGCAAAAAGAGCTGGCGGGCGCTGGCCACGTCGCTGGG 240  
Qy 241 GCGCGGCGGAGCGTGGAGTGGCCAGCTGCAACCGGGCGGGCTCGGCTTTGTTGCGCA 300  
Db 241 GCGCGGCGGAGCGTGGAGTGGCCAGCTGCAACCGGGCGGGCTCGGCTTTGTTGCGCA 300  
Qy 301 GTAAAGGAGGAGGCTGTCTCAGCTGCAGAGGGGTCTATCCCTGTCTTCAAGCAGTGCCTC 360  
Db 301 GTAAAGGAGGAGGCTGTCTCAGCTGCAGAGGGGTCTATCCCTGTCTTCAAGCAGTGCCTC 360

QY 361 TTCCAGCTCCCATGGGACCCGAAAGCCAGCTCCGGATGGAAACCTGGACGTGAAG 420  
DB 361 TTCCAGCTCCCATGGGACCCGAAAGCCAGCTCCGGATGGAAACCTGGACGTGAAG 420  
QY 421 GAGGAATGACAGACCAAGATCTTCCAGGSCCACTCCCAAGAGAGACGGGGTGGAACTG 480  
DB 421 GAGGAATGACAGACCAAGATCTTCCAGGSCCACTCCCAAGAGAGACGGGGTGGAACTG 480  
QY 481 CTTGGCAGCCCGTGGAGACACATCTCTCTCCCAACACGCTAAATTTCAACGGAGCG 540  
DB 481 CTTGGCAGCCCGTGGAGACACATCTCTCTCCCAACACGCTAAATTTCAACGGAGCG 540  
QY 541 CATCTGAAGAGAGACGCTGGTGGCCCCCAGAGATCAACATTTCTCTGGATCAGAGTGAG 600  
DB 541 CATCTGAAGAGAGACGCTGGTGGCCCCCAGAGATCAACATTTCTCTGGATCAGAGTGAG 600  
QY 601 GGGTCCCTCTCTCCGATGACTCTTTGGATACCCCTGATGACTGGATTAACCTGGAT 660  
DB 601 GGGTCCCTCTCTCCGATGACTCTTTGGATACCCCTGATGACTGGATTAACCTGGAT 660  
QY 661 GACATCGAGACCCCGATGAGACCCGACTCGCTGGAGTTCTGGGGAATGGCAACGAATG 720  
DB 661 GACATCGAGACCCCGATGAGACCCGACTCGCTGGAGTTCTGGGGAATGGCAACGAATG 720  
QY 721 GAGTGGGAAGAAGACACCCCGTGGCCAACCGCAAGAACATGCCCCGGGACACGCGGAT 780  
DB 721 GAGTGGGAAGAAGACACCCCGTGGCCAACCGCAAGAACATGCCCCGGGACACGCGGAT 780  
QY 781 CTATTTGGGAGCGGACGAGAGGAGCGGAGCGGCGGCCCAACGGGGCCCTGTGGCGGACA 840  
DB 781 CTATTTGGGAGCGGACGAGAGGAGCGGAGCGGCGGCCCAACGGGGCCCTGTGGCGGACA 840  
QY 841 GTGATCATCGGGAGCAAGAGACACCGTATAGACCTGACATGATCCGGCTTTACATGAAA 900  
DB 841 GTGATCATCGGGAGCAAGAGACACCGTATAGACCTGACATGATCCGGCTTTACATGAAA 900  
QY 901 GTGGTCAACCAACGGAGGTAATAAGCGAAGGCTCAACGCCATCATCTCTTCGAGCC 960  
DB 901 GTGGTCAACCAACGGAGGTAATAAGCGAAGGCTCAACGCCATCATCTCTTCGAGCC 960  
QY 961 TSCCTTCTCCAGACAGGAGCTCCCGGACTACCACTACATCATGAGAGACCTCTTCCTG 1020  
DB 961 TSCCTTCTCCAGACAGGAGCTCCCGGACTACCACTACATCATGAGAGACCTCTTCCTG 1020  
QY 1021 TAGCTCATCAGCAGCTTAGAGCTCTCTGGTGGTGAAGCTACATGATCGTGTACCTGAAC 1080  
DB 1021 TAGCTCATCAGCAGCTTAGAGCTCTCTGGTGGTGAAGCTACATGATCGTGTACCTGAAC 1080  
QY 1081 GGTGCCACGCGCGGAGGATGCTGGAATCGGCTGGCTGAAGAAGTGTACAGATG 1140  
DB 1081 GGTGCCACGCGCGGAGGATGCTGGAATCGGCTGGCTGAAGAAGTGTACAGATG 1140  
QY 1141 ATCGACCGAGGTTGGGAAACCTGAAAGTCTTGTATCATGTCACCCCTCTGGGTTTC 1200  
DB 1141 ATCGACCGAGGTTGGGAAACCTGAAAGTCTTGTATCATGTCACCCCTCTGGGTTTC 1200  
QY 1201 ATTCCGACTGTCTGGCCATCTCTCGCCCTTTTCATCAGCGGTCAAGTTTCATCAACAGATC 1260  
DB 1201 ATTCCGACTGTCTGGCCATCTCTCGCCCTTTTCATCAGCGGTCAAGTTTCATCAACAGATC 1260  
QY 1261 CAGTACGTGCAGCTTGAAGACCTGGAGCAACTCATCCCTATGGAAACGCTCCAGATC 1320  
DB 1261 CAGTACGTGCAGCTTGAAGACCTGGAGCAACTCATCCCTATGGAAACGCTCCAGATC 1320  
QY 1321 CCAGACTCGTCTGCAATACGAGAGGAAAGACTGAAGCCAGGAGGAGAGAGCGGAGG 1380  
DB 1321 CCAGACTCGTCTGCAATACGAGAGGAAAGACTGAAGCCAGGAGGAGAGAGCGGAGG 1380  
QY 1381 CCCCAGCCGAGTTTGTCTGCTGCGCAGGTCTGAAGAGAGGAGAGGTGGCACAGTGGAA 1440  
DB 1381 CCCCAGCCGAGTTTGTCTGCTGCGCAGGTCTGAAGAGAGGAGAGGTGGCACAGTGGAA 1440  
QY 1441 AACAGGCTCTGCTCTCAGAAATCAGGAAACAGCATGTCTCTGAGGCGAGCTGAGC 1500

DB 1441 AACAGGCTCTGCTCTCAGAAATCAGGAAACAGCATGTCTGAGGCGACGTGAGC 1500  
QY 1501 ATAACAAAGGACATGGAAGAGATTCAGATGCGCAAGAAACCTCTGTCTGAGCGCCACTG 1560  
DB 1501 ATAACAAAGGACATGGAAGAGATTCAGATGCGCAAGAAACCTCTGTCTGAGCGCCACTG 1560  
QY 1561 GCCCCAGATCTCATCTCTGCTCATCTCTGAGTCCCAATCTTCCAAAGGTCGCAAGCCCTCC 1620  
DB 1561 GCCCCAGATCTCATCTCTGCTCATCTCTGAGTCCCAATCTTCCAAAGGTCGCAAGCCCTCC 1620  
QY 1621 GTTCATCTCTGAAACCCAGCATCTCTTTTCACTGCTTCAAAACATTTGATTTTTTTTTT 1680  
DB 1621 GTTCATCTCTGAAACCCAGCATCTCTTTTCACTGCTTCAAAACATTTGATTTTTTTTTT 1680  
QY 1681 TAAAGATCGAGTATTTGTGCGTTCCAGAAAGGGCCAGCTCTGAGCCCTCACCCCTTCC 1740  
DB 1681 TAAAGATCGAGTATTTGTGCGTTCCAGAAAGGGCCAGCTCTGAGCCCTCACCCCTTCC 1740  
QY 1741 ACATCTCAAGAACTCTCAGCCGAGGAGGCAAGAGCGGAGGGTGGCCCGCTGGCGTC 1800  
DB 1741 ACATCTCAAGAACTCTCAGCCGAGGAGGCAAGAGCGGAGGGTGGCCCGCTGGCGTC 1800  
QY 1801 GGTGGCTCTCCCTCTCTGCTCGAGCCCTGTGGTCTGAGTGGATACAAGATTTCAAGACC 1860  
DB 1801 GGTGGCTCTCCCTCTCTGCTCGAGCCCTGTGGTCTGAGTGGATACAAGATTTCAAGACC 1860  
QY 1861 CTTCTCTTGTCTGTCACCCGCTCCAGGTTGAGGCAAGACCCACCGCCACCCCGGCT 1920  
DB 1861 CTTCTCTTGTCTGTCACCCGCTCCAGGTTGAGGCAAGACCCACCGCCACCCCGGCT 1920  
QY 1921 GGGTCTGCTCTCTTCTGCTGCTTCCCTCCAGAAATCGGCTCAGACCTAGAAAGCTCA 1980  
DB 1921 GGGTCTGCTCTCTTCTGCTGCTTCCCTCCAGAAATCGGCTCAGACCTAGAAAGCTCA 1980  
QY 1981 ACCCCCTTATGAGGSCACGCTCTGGGTAGCTCTCGACCTCGACCTTATGTCCAAAT 2040  
DB 1981 ACCCCCTTATGAGGSCACGCTCTGGGTAGCTCTCGACCTCGACCTTATGTCCAAAT 2040  
QY 2041 TCACACCCATGGTTTTCATTTGACCCGCGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 2100  
DB 2041 TCACACCCATGGTTTTCATTTGACCCGCGCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 2100  
QY 2101 TTTGAGAGGATCAGAGCCCATTTGCAAGAAAGAGCGCTGCGCAACCTCTTGTCTCGG 2160  
DB 2101 TTTGAGAGGATCAGAGCCCATTTGCAAGAAAGAGCGCTGCGCAACCTCTTGTCTCGG 2160  
QY 2161 ATTGCAAAATGACACCCAGTAATCTAGAACATTTCTCAAGCCCTTTAACTCAGATGCA 2220  
DB 2161 ATTGCAAAATGACACCCAGTAATCTAGAACATTTCTCAAGCCCTTTAACTCAGATGCA 2220  
QY 2221 AGCCACCGGGCAACCCCGTCAATCTCCACCAAGGAATGAGATATGTGGACCTCACT 2280  
DB 2221 AGCCACCGGGCAACCCCGTCAATCTCCACCAAGGAATGAGATATGTGGACCTCACT 2280  
QY 2281 GCTCCCCCAACCCAGCGTCAAGTGGGACACGCGCAACGCTGTTCCGGGTTGGAAACGAG 2340  
DB 2281 GCTCCCCCAACCCAGCGTCAAGTGGGACACGCGCAACGCTGTTCCGGGTTGGAAACGAG 2340  
QY 2341 AGGCTCAGAAATCTGCTCTGAAATAGGAGACCTAGCAAGAGGAAGATACAGGGTATCGG 2400  
DB 2341 AGGCTCAGAAATCTGCTCTGAAATAGGAGACCTAGCAAGAGGAAGATACAGGGTATCGG 2400  
QY 2401 GGGTTGAGTCTTTCAGAGTCAATTCGGGAAGTAATCCAGTGGCTGGCGCGGAGCCAC 2460  
DB 2401 GGGTTGAGTCTTTCAGAGTCAATTCGGGAAGTAATCCAGTGGCTGGCGCGGAGCCAC 2460  
QY 2461 CTGCATTCAGAGCTTGGACCGGGTTCTTGTTCGGGAGGCAAAATTTCCCTAGGAAAA 2520  
DB 2461 CTGCATTCAGAGCTTGGACCGGGTTCTTGTTCGGGAGGCAAAATTTCCCTAGGAAAA 2520  
QY 2521 GAAGACAGCTTTTCTAAATGTGGTCCAAATGCGGATCACTGGTCTCAGATGGACTCTAGAG 2580



Db 2521 GAAGACAGACTTTCTAATGTGGTCCAAATGCGGATCACTGGTCAAGTGAAG 2580  
Qy 2581 CACTGAGCTCCCTGTCTCTGGAAGTATTAAAGAAAGCTGGGCCAGGCAAGTGGCTCA 2640  
Db 2581 CACTGAGCTCCCTGTCTCTGGAAGTATTAAAGAAAGCTGGGCCAGGCAAGTGGCTCA 2640  
Qy 2641 CGCCTGTAAATCCAGACTTTGGAGGCCCGAGGCGGCGGATCACTGAGGTGAGAGTTT 2700  
Db 2641 CGCCTGTAAATCCAGACTTTGGAGGCCCGAGGCGGCGGATCACTGAGGTGAGAGTTT 2700  
Qy 2701 GAGAACAGCTGGCCAAATGTTGAACCTCATCTCTCTAAATAATACAAATATTAGCCA 2760  
Db 2701 GAGAACAGCTGGCCAAATGTTGAACCTCATCTCTCTAAATAATACAAATATTAGCCA 2760  
Qy 2761 GCGGTGGTGGCAGGTGCTGTAAATCCAGCTACTTTGGAGGCTGAGGCAATGAGATCACT 2820  
Db 2761 GCGGTGGTGGCAGGTGCTGTAAATCCAGCTACTTTGGAGGCTGAGGCAATGAGATCACT 2820  
Qy 2821 TAAACTGAGAGCGAGAGTTACAGTGAGCCAAAGATCGTGCCAATTCAGGCTGGG 2880  
Db 2821 TAAACTGAGAGCGAGAGTTACAGTGAGCCAAAGATCGTGCCAATTCAGGCTGGG 2880  
Qy 2881 CGACAGAGCAAGACTCTGTCTCAAAAAAAGAAAAA 2918  
Db 2881 CGACAGAGCAAGACTCTGTCTCAAAAAAAGAAAAA 2918

## RESULT 2

US-10-699-941-8  
; Sequence 8, Application US/10699941  
; Publication No. US20040146900A1  
; GENERAL INFORMATION:  
; APPLICANT: Burmeister, Margit  
; TITLE OF INVENTION: Ataxia Associated Gene and Protein  
; FILE REFERENCE: UM-08441  
; CURRENT APPLICATION NUMBER: US/10/699,941  
; CURRENT FILING DATE: 2003-11-03  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 8  
; LENGTH: 2918  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-699-941-8

Query Match 99.9%; Score 2916.4; DB 7; Length 2918;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2917; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GCCGAGCCTCTGCCAGCCTCTGAGCTGGGAAGAGCAGCTACTCTCGGAGGCGAGGCGCA 60  
Db 1 GCCGAGCCTCTGCCAGCCTCTGAGCTGGGAAGAGCAGCTACTCTCGGAGGCGAGGCGCA 60  
Qy 61 GCGGCGCGGCGATGAGAGGCGGCGCAGCGCAGCCCGCTGGGAGCCACCGCTAAC 120  
Db 61 GCGGCGCGGCGATGAGAGGCGGCGCAGCGCAGCCCGCTGGGAGCCACCGCTAAC 120  
Qy 121 CCTGCACCCACCCACCTTCACAAAGAGCTGGCGGCGCTGGCCAGCTGGCCTGGG 180  
Db 121 CCTGCACCCACCCACCTTCACAAAGAGCTGGCGGCGCTGGCCAGCTGGCCTGGG 180  
Qy 181 TGACCTTCTCTGGATGACAGAAATCCCGCTTCGAGCATCTCTTCTCTTAGGCTCTGAA 240  
Db 181 TGACCTTCTCTGGATGACAGAAATCCCGCTTCGAGCATCTCTTCTCTTAGGCTCTGAA 240  
Qy 241 GCGCGGCGGAGGCTGAGCGATGCCAGCTGCACCGGCGAGGCTCGCCTTTGTTGCCA 300  
Db 241 GCGCGGCGGAGGCTGAGCGATGCCAGCTGCACCGGCGAGGCTCGCCTTTGTTGCCA 300  
Qy 301 GTAAGGAGAGGCTGTCTAGCTGCGAGGGGTATCCCTTGTCTTAAGCAGTGCCTC 360  
Db 301 GTAAGGAGAGGCTGTCTAGCTGCGAGGGGTATCCCTTGTCTTAAGCAGTGCCTC 360

Qy 361 TTCCAGCTCCCATGGGACCAACGAGCCACGCTCCCGATCGAATAAAGCTGACCTGAAG 420  
Db 361 TTCCAGCTCCCATGGGACCAACGAGCCACGCTCCCGATCGAATAAAGCTGACCTGAAG 420  
Qy 421 GAGGAATGGCAGGACGAAGATCTTCCAGGCGCACTCCCAAGAGAGACGCGGGGTGAAGTG 480  
Db 421 GAGGAATGGCAGGACGAAGATCTTCCAGGCGCACTCCCAAGAGAGACGCGGGGTGAAGTG 480  
Qy 481 CTTGGCAGCCCGGTGGAGAGACATCTCTCTCCCTCCAAACGCTAAATTTCAACGAGCG 540  
Db 481 CTTGGCAGCCCGGTGGAGAGACATCTCTCTCCCTCCAAACGCTAAATTTCAACGAGCG 540  
Qy 541 CATCGTAAGAGAGAGACGCTGGTGGCCCCAGAGATCAACATTTCTCTGATCAGAGTGA 600  
Db 541 CATCGTAAGAGAGAGACGCTGGTGGCCCCAGAGATCAACATTTCTCTGATCAGAGTGA 600  
Qy 601 GGGTCCCTGTCTGTCGATGACTTTCTTGGATATACCTCTGATGATTAACCTGGAT 660  
Db 601 GGGTCCCTGTCTGTCGATGACTTTCTTGGATATACCTCTGATGATTAACCTGGAT 660  
Qy 661 GACATCGAGACCCCGATGAGACCGACTCGCTGGAGTTCTTGGGGAATGGCAACGACTG 720  
Db 661 GACATCGAGACCCCGATGAGACCGACTCGCTGGAGTTCTTGGGGAATGGCAACGACTG 720  
Qy 721 GAGTGGGAAGACGACACCCCGTGGCCACCGCCAAAGAACATGCCCGGGGACAGCGCGAT 780  
Db 721 GAGTGGGAAGACGACACCCCGTGGCCACCGCCAAAGAACATGCCCGGGGACAGCGCGAT 780  
Qy 781 CTATTTGGGAGCGGCA CGACGAGAGCGGCGCGCCAAACGCGGCGCTGTGGCGGACA 840  
Db 781 CTATTTGGGAGCGGCA CGACGAGAGCGGCGCGCCAAACGCGGCGCTGTGGCGGACA 840  
Qy 841 GTGATCATCGGGGAGCAAGACCGCTATAGACCTGACATGATTCGGGCTTACATGAAA 900  
Db 841 GTGATCATCGGGGAGCAAGACCGCTATAGACCTGACATGATTCGGGCTTACATGAAA 900  
Qy 901 GTGGTCAACCACGAGGCTACTAGCGGAGGCGCTCAAGCGCATCATCTCTTCGCGCC 960  
Db 901 GTGGTCAACCACGAGGCTACTAGCGGAGGCGCTCAAGCGCATCATCTCTTCGCGCC 960  
Qy 961 TGCTTCTTCAGACAGCGCTCCCGACTACCACTACATCATGAGAGAACCTTCTCTG 1020  
Db 961 TGCTTCTTCAGACAGCGCTCCCGACTACCACTACATCATGAGAGAACCTTCTCTG 1020  
Qy 1021 TACGTATCAGCAGCTTAGAGCTCTCTGGTGGCTGAGGACTACATGATCGTGTACCTGAAC 1080  
Db 1021 TACGTATCAGCAGCTTAGAGCTCTCTGGTGGCTGAGGACTACATGATCGTGTACCTGAAC 1080  
Qy 1081 GGTGCCAGCCCGCGGAGGATGCTGGAAATCGGCTGGAAGAGTGTACAGATG 1140  
Db 1081 GGTGCCAGCCCGCGGAGGATGCTGGAAATCGGCTGGAAGAGTGTACAGATG 1140  
Qy 1141 ATCGACCGGAGTTTCGGGAAAAAAGCTGAAAGTCTTGTGATCATCTGCTCCACCCCTCGTGGTTC 1200  
Db 1141 ATCGACCGGAGTTTCGGGAAAAAAGCTGAAAGTCTTGTGATCATCTGCTCCACCCCTCGTGGTTC 1200  
Qy 1201 ATTGCGACTGTGCTGGCCATCTCTGCGCCCTTTTCAATCAGCGTCAAGTTTCAACAAGATC 1260  
Db 1201 ATTGCGACTGTGCTGGCCATCTCTGCGCCCTTTTCAATCAGCGTCAAGTTTCAACAAGATC 1260  
Qy 1261 CAGTACGTGCAACGCTTGGAGACCTGGAGCACTCATCTCCCTATGGAACAGCTCCAGATC 1320  
Db 1261 CAGTACGTGCAACGCTTGGAGACCTGGAGCACTCATCTCCCTATGGAACAGCTCCAGATC 1320  
Qy 1321 CCAGACTGCGTCTCTCAATACGAAGAGAAAGACTGAAGGCCAGGAGGAGGCGGAGG 1380  
Db 1321 CCAGACTGCGTCTCTCAATACGAAGAGAAAGACTGAAGGCCAGGAGGAGGCGGAGG 1380  
Qy 1381 CCCAGCGGAGGTTTGTGCTGCCAGGTCTGAAGAGAGCCAGAGGTGGCCAGGTGGA 1440  
Db 1381 CCCAGCGGAGGTTTGTGCTGCCAGGTCTGAAGAGAGCCAGAGGTGGCCAGGTGGA 1440  
Qy 1441 AACAGGTCTGCTCTGCTCTCAGAGATCAGGAACCAAGCATGTCTCTGAGGCGAGCTGAGC 1500

```
Db 1441 AACAGCTCTGCTGCTCTCAGAAATCAGGAAACAAAGCATGTCTTGAGGCGACGTGAGC 1500
Qy 1501 ATAACAAAGACATGGAAGAAGATTCAGATGCCAGAAACCTCTGTGTCAGAGCGCCACTG 1560
Db 1501 ATAACAAAGACATGGAAGAAGATTCAGATGCCAGAAACCTCTGTGTCAGAGCGCCACTG 1560
Qy 1561 GCGCCAGATCTCATCTGCTGCTCATCTGATGCTCCAAATCTTCAAGGGTCCAGCCCTCC 1620
Db 1561 GCGCCAGATCTCATCTGCTGCTCATCTGATGCTCCAAATCTTCAAGGGTCCAGCCCTCC 1620
Qy 1621 GTTCATCTCTGAAACCCAGCATCTTTTTCAGTCTGCTTGAACATGTGATTTTTTTTTT 1680
Db 1621 GTTCATCTCTGAAACCCAGCATCTTTTTCAGTCTGCTTGAACATGTGATTTTTTTTTT 1680
Qy 1681 TAAAGATGATGATTTTGTGCTTCCAGAAAGGGCCAGCTCTGAGCCCTTCACTCTCC 1740
Db 1681 TAAAGATGATGATTTTGTGCTTCCAGAAAGGGCCAGCTCTGAGCCCTTCACTCTCC 1740
Qy 1741 AACTACAGAACTCTCAGCCGAGGAAGCAAGAGCGCAGGGGTGGCCGCTGGCGTC 1800
Db 1741 AACTACAGAACTCTCAGCCGAGGAAGCAAGAGCGCAGGGGTGGCCGCTGGCGTC 1800
Qy 1801 GGTGGCTCCGCTCTGCTCGAGCCCTGTGCTGAGCTGGATCAGATTCAGACC 1860
Db 1801 GGTGGCTCCGCTCTGCTCGAGCCCTGTGCTGAGCTGGATCAGATTCAGACC 1860
Qy 1861 CTCTCTGCTGCTGCTACCCGCTTCCAGGTTGGAGCCACAGACACCCAGCCACCCGGCT 1920
Db 1861 CTCTCTGCTGCTGCTACCCGCTTCCAGGTTGGAGCCACAGACACCCAGCCACCCGGCT 1920
Qy 1921 GGGTCTGCTGCTCTTCTGCTGCTTCCCTCAGAAATGGGCTCAGACCTAGAAAGTCA 1980
Db 1921 GGGTCTGCTGCTCTTCTGCTGCTTCCCTCAGAAATGGGCTCAGACCTAGAAAGTCA 1980
Qy 1981 ACCCCCTATGAGGCGCAGTCTGGGTAGCTCTGACTCGACCTCGACCTTATGTCAAAT 2040
Db 1981 ACCCCCTATGAGGCGCAGTCTGGGTAGCTCTGACTCGACCTCGACCTTATGTCAAAT 2040
Qy 2041 TCACACCCATGTTTTCATTTGACCCGCCCTTCTGCTCATATGACACCCAGCTCC 2100
Db 2041 TCACACCCATGTTTTCATTTGACCCGCCCTTCTGCTCATATGACACCCAGCTCC 2100
Qy 2101 TTTGAGAGATCAGAGCCCATTTGCAACAAGAGCGCTGCAACCATCTTTGCTCGG 2160
Db 2101 TTTGAGAGATCAGAGCCCATTTGCAACAAGAGCGCTGCAACCATCTTTGCTCGG 2160
Qy 2161 ATTGCAAAATGACCCCAAGTAACTAGAACATTTCTAAGCCCTTTAACTCAGATGCA 2220
Db 2161 ATTGCAAAATGACCCCAAGTAACTAGAACATTTCTAAGCCCTTTAACTCAGATGCA 2220
Qy 2221 AGCCACCGGCAAAACCCCGTCAATACCTCCACCAAGGAATGAGATATGTGACCTCACT 2280
Db 2221 AGCCACCGGCAAAACCCCGTCAATACCTCCACCAAGGAATGAGATATGTGACCTCACT 2280
Qy 2281 GCTCCCCCAACCCAGCGTCAAGGTGGGACAACGCAACGCTGTTCCGGGTGGAAACAGCAG 2340
Db 2281 GCTCCCCCAACCCAGCGTCAAGGTGGGACAACGCAACGCTGTTCCGGGTGGAAACAGCAG 2340
Qy 2341 AGGCTCAGAAATCTGCTGTAATAGGACAGCTTAGCAGAGGAGATACAGGGTATCGG 2400
Db 2341 AGGCTCAGAAATCTGCTGTAATAGGACAGCTTAGCAGAGGAGATACAGGGTATCGG 2400
Qy 2401 GGGTTGAGTGTTCAGAAATCATTGCGGAGAGTAATTCAGTGCCTGCGCCGAGCCAC 2460
Db 2401 GGGTTGAGTGTTCAGAAATCATTGCGGAGAGTAATTCAGTGCCTGCGCCGAGCCAC 2460
Qy 2461 CTGCAATCAAGCTTGGACAGCGGTTCTTGTTCGGAGGCAAAATTCCTTAGGAAAA 2520
Db 2461 CTGCAATCAAGCTTGGACAGCGGTTCTTGTTCGGAGGCAAAATTCCTTAGGAAAA 2520
Qy 2521 GAAGACAGACTTTCTTAATGTGTTCCAAATGCGGATCACTGTCAGATGATCTTAGAG 2580
```

```
Db 2521 GAAGACAGACTTTCTTAATGTGTTCCAAATGCGGATCACTGTCAGATGGAATCTTAGAG 2580
Qy 2581 CACTGAGCTCCCTGCTCTCTGGAAGTATTTAAGAAAAGGCTGGGCGCAGGCAACGATGCTCA 2640
Db 2581 CACTGAGCTCCCTGCTCTCTGGAAGTATTTAAGAAAAGGCTGGGCGCAGGCAACGATGCTCA 2640
Qy 2641 CCGCTGTAATCCAGACTTTTGGAGGCGGAGCGGATCCTCATCTACTATAAAATACAAAATTAGCCA 2700
Db 2641 CCGCTGTAATCCAGACTTTTGGAGGCGGAGCGGATCCTCATCTACTATAAAATACAAAATTAGCCA 2700
Qy 2701 GAGACAGCTGGGCGCAACATGTTGAAACCTCATCTACTATAAAATACAAAATTAGCCA 2760
Db 2701 GAGACAGCTGGGCGCAACATGTTGAAACCTCATCTACTATAAAATACAAAATTAGCCA 2760
Qy 2761 GCGCTGGTGGCAGGCTGCTGTAATCCAGCTACTTTGGAGGCTGAGGCAATGAGATCACT 2820
Db 2761 GCGCTGGTGGCAGGCTGCTGTAATCCAGCTACTTTGGAGGCTGAGGCAATGAGATCACT 2820
Qy 2821 TAAACCTGAGAGGCGAGAGTTTACAGTGAGCAGATCGTCCACTGCAATCCAGCTGG 2880
Db 2821 TAAACCTGAGAGGCGAGAGTTTACAGTGAGCAGATCGTCCACTGCAATCCAGCTGG 2880
Qy 2881 CGACAGCAAGACTCTGCTCTCAAAAAA 2918
Db 2881 CGACAGCAAGACTCTGCTCTCAAAAAA 2918

RESULT 3
US-10-104-047-997
; Sequence 997, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 997
; LENGTH: 2759
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-997

Query Match 89.0%; Score 2598.2; DB 6; Length 2759;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 2756; Conservative 0; Mismatches 3; Indels 136; Gaps 2;

Qy 8 CTCTGCCAGCCCTGAGCTGGGAAGAGCAGCTACCTCGAGGAGCGGCGCGGCGGCGG 67
Db 1 CTCTGCCAGCCCTGAGCTGGGAAGAGCAGCTACCTCGAGGAGCGGCGCGGCGGCGG 60
Qy 68 GCGCATGAGAGGCGGCGCAGCGCCGCGCTGGGAGGCCACCCGCTAACCCCTGCAC 127
Db 61 GCGCATGAGAGGCGGCGCAGCGCCGCGCTGGGAGGCCACCCGCTAACCCCTGCAC 120
Qy 128 CCCACCCACCCCTGCACAAAAGAGCTGGCGGCGCTGGCCACGTCGCTTGGGTGACCTT 187
Db 121 CCCACCCA----- 128
Qy 188 CCTCGATGCAGATCCCGCTCGAGCATCTCTTCTCTAGGCTCTGAAGGCCCGG 247
Db 129 -----CCCTTCGAGCATCTCTTCTCTAGGCTCTGAAGGCCCGG 170
Qy 248 GGAGCGTGAAGCGATGCCAGCTGCACCCGCGGCGGCTCGCTTGTGTTGTCAGTAAGA 307
Db 171 GGAGCGTGAAGCGATGCCAGCTGCACCCGCGGCGGCTCGCTTGTGTTGTCAGTAAGA 230
Qy 308 GGAGGCGTGTCTCAGCTGCAAGAGGGGTCACTCCCTGCTTCAAGCAGTGCCTTTCCAG 367
```

Db 231 GGAGAGGCTGTCTCAGCTGCAGAGGGGTCACTCCCTGCTTCAAGCCAGTGCCTCTTCCGAG 290  
Qy 368 CTCCCATGGGGAACAACGGAAGCAACGCTCCGGATGGAACAGTGGACGCTGAAGGAGAAAT 427  
Db 291 CTCCCATGGGGAACAACGGAAGCAACGCTCCGGATGGAACAGTGGACGCTGAAGGAGAAAT 350  
Qy 428 GGAGAGGAGAGATCTTCCAGGCACTCCAGAGAGAGAGGGGGTGGAACTGCTTGGCA 487  
Db 351 GGC----- 353  
Qy 488 GCCGGTGGGAAGACACATCTCTCTCCCAACACGCTTAAATTTCAACGGAGCGCATCGTA 547  
Db 354 -----AAGACACATCTCTCTCCCAACACGCTTAAATTTCAACGGAGCGCATCGTA 404  
Qy 548 AAGAGAGAGCGTGGTGGGCCCAAGAGATCAACATTTCTCTGGATCAGAGTGAAGGGTCCC 607  
Db 405 AAGAGAGAGCGTGGTGGGCCCAAGAGATCAACATTTCTCTGGATCAGAGTGAAGGGTCCC 464  
Qy 608 TGTGTCCGATGACTTCTTGGATACCCCTGTGATGACCTGGATATTAACGTGGATGACATCG 667  
Db 465 TGTGTCCGATGACTTCTTGGATACCCCTGTGATGACCTGGATATTAACGTGGATGACATCG 524  
Qy 668 AGACCCCCCATGAGACCGACTCGCTGGAGTTCTCTGGGGAATGSCAACGAACTGGAGTGG 727  
Db 525 AGACCCCCCATGAGACCGACTCGCTGGAGTTCTCTGGGGAATGSCAACGAACTGGAGTGG 584  
Qy 728 AAGAGACACCCCGTGGGACACCGCCAGAAACATATGCCCCGGGACACGCGGATCTATTG 787  
Db 585 AAGAGACACCCCGTGGGACACCGCCAGAAACATATGCCCCGGGACACGCGGATCTATTG 644  
Qy 788 GGGAGCGACGACGAGAGAGCGGAGCGCCGCAACCGGGCGCTGTGGCGGACAGTGAATCA 847  
Db 645 GGGAGCGACGACGAGAGAGCGGAGCGCCGCAACCGGGCGCTGTGGCGGACAGTGAATCA 704  
Qy 848 TCGGGAGCAGAGACCGGTATAGACTGCAATGATCCGGCTTACATGAAGTGGTCA 907  
Db 705 TCGGGAGCAGAGACCGGTATAGACTGCAATGATCCGGCTTACATGAAGTGGTCA 764  
Qy 908 CCACGGAGGTACTACGGCGAAGGCTCAACGCCATCATCGTCTTTCGAGCGCTGCTCC 967  
Db 765 CCACGGAGGTACTACGGCGAAGGCTCAACGCCATCATCGTCTTTCGAGCGCTGCTCC 824  
Qy 968 TTCAGACAGACGCTCCCGATACCACTACATCATGAGAACTCTTCTGTAAGTCA 1027  
Db 825 TTCAGACAGACGCTCCCGATACCACTACATCATGAGAACTCTTCTGTAAGTCA 884  
Qy 1028 TCAGACGCTTACAGCTCTGCTGGTGGAGCTACATGATCGTGTACCTGAAGGTTGCA 1087  
Db 885 TCAGACGCTTACAGCTCTGCTGGTGGAGCTACATGATCGTGTACCTGAAGGTTGCA 944  
Qy 1088 CGCCCGCGGAGGATCGCTGGAACTCGGCTGGAAGGTGCTACCAAGATGATCGAC 1147  
Db 945 CGCCCGCGGAGGATCGCTGGAACTCGGCTGGAAGGTGCTACCAAGATGATCGAC 1004  
Qy 1148 GGAGGTGGGGAAGAACTGAGTCTTGATCATCGTCCACCCTCGTGGTTCAATTCGGA 1207  
Db 1005 GGAGGTGGGGAAGAACTGAGTCTTGATCATCGTCCACCCTCGTGGTTCAATTCGGA 1064  
Qy 1208 CTGTGTGGCCATCTCTCGCCCTTTCATCAGAGTCAAGTTCAATCAAGATTCAGTAG 1267  
Db 1065 CTGTGTGGCCATCTCTCGCCCTTTCATCAGAGTCAAGTTCAATCAAGATTCAGTAG 1124  
Qy 1268 TCACAGCTTGGAGACCTGGAGCAACTCATCCCTATGGAACACGCTCCAGATCCAGACT 1327  
Db 1125 TCACAGCTTGGAGACCTGGAGCAACTCATCCCTATGGAACACGCTCCAGATCCAGACT 1184  
Qy 1328 GCGTCTGCAATACGAGAGGAAAGACTGAAGGCCAGGAGGAGCGCGAGGCCCGCAGC 1387  
Db 1185 GCGTCTGCAATACGAGAGGAAAGACTGAAGGCCAGGAGGAGCGCGAGGCCCGCAGC 1244  
Qy 1388 CGGAGTTGTGCTGCCAGGTCTGAAGAGAGCCAGAGTGGCAACAGTGGAAACAGGT 1447  
Db 1245 CGGAGTTGTGCTGCCAGGTCTGAAGAGAGCCAGAGTGGCACACAGTGGAAACAGGT 1304

Qy 1448 CTGCTCTGCTCTCAGAGATCAGGAAACAAGCATGTCTCTAGGCGCAGCTGAGCAATAACA 1507  
Db 1305 CTGCTCTGCTCTCAGAGATCAGGAAACAAGCATGTCTCTAGGCGCAGCTGAGCAATAACA 1364  
Qy 1508 AGGACATGGAAGAGATTCAGATGCCAGAAACCTCTGTGTCAGACGCCCACTGGGCCAG 1567  
Db 1365 AGGACATGGAAGAGATTCAGATGCCAGAAACCTCTGTGTCAGACGCCCACTGGGCCAG 1424  
Qy 1568 ATCTCATCTGCTCATCTGAGTCCCAATCTTCAAGGGTGCAGGCCCTCGGTTCTC 1627  
Db 1425 ATCTCATCTGCTCATCTGAGTCCCAATCTTCAAGGGTGCAGGCCCTCGGTTCTC 1484  
Qy 1628 TCTGAAACCCAGCATCTCTTTTTCAGCTGTCTGAAACATTTATTTTTTTTTTAAACGAT 1687  
Db 1485 TCTGAAACCCAGCATCTCTTTTTCAGCTGTCTGAAACATTTATTTTTTTTTTAAACGAT 1544  
Qy 1688 GCAGTATTTGTGCTTTCAGAAAGGGCCAGCTCTGAGGCCCTCAACCTTCCACATCA 1747  
Db 1545 GCAGTATTTGTGCTTTCAGAAAGGGCCAGCTCTGAGGCCCTCAACCTTCCACATCA 1604  
Qy 1748 CGAACTCTCAGCGGAGGAGCAAGAGCGCAGGGGGTGGGCCGCTGGCGTGGTGGCC 1807  
Db 1605 CGAACTCTCAGCGGAGGAGCAAGAGCGCAGGGGGTGGGCCGCTGGCGTGGTGGCC 1664  
Qy 1808 TCCGCTCTGCTCGCAGGCCCTGTGCTCAGAGCTGATATCAAGATTTCAAGACCTTCTCT 1867  
Db 1665 TCCGCTCTGCTCGCAGGCCCTGTGCTCAGAGCTGATATCAAGATTTCAAGACCTTCTCT 1724  
Qy 1868 TGTGCTCACCGCTTCCAGGTTGGAGCCACAGACCCACCGCCACCGCTGGGCTGCTG 1927  
Db 1725 TGTGCTCACCGCTTCCAGGTTGGAGCCACAGACCCACCGCCACCGCTGGGCTGCTG 1784  
Qy 1928 GGTCTTCTGCTGCTTCTCCCTCCAGAAATGGGCTCAGACCTAGAAAGCTCAACCCGCC 1987  
Db 1785 GGTCTTCTGCTGCTTCTCCCTCCAGAAATGGGCTCAGACCTAGAAAGCTCAACCCGCC 1844  
Qy 1988 TATGAGGCGCAGCTCTCGGCTAGCTCTGACCTCCGACCTTATGTCCAAATTTTCACACC 2047  
Db 1845 TATGAGGCGCAGCTCTCGGCTAGCTCTGACCTCCGACCTTATGTCCAAATTTTCACACC 1904  
Qy 2048 CATGTTTTTTCATTTGACCCGCCCTTCTGCTCATATGACACCCAGCTCTTTTGAGA 2107  
Db 1905 CATGTTTTTTCATTTGACCCGCCCTTCTGCTCATATGACACCCAGCTCTTTTGAGA 1964  
Qy 2108 GGATCAGAGCCCATTTGACAAAGAGAGCGCTGCCAACCATCTTGTCTCTCGATGCA 2167  
Db 1965 GGATCAGAGCCCATTTGACAAAGAGAGCGCTGCCAACCATCTTGTCTCTCGATGCA 2024  
Qy 2168 AATGACACCCCAAGTAATCTAGAACATTTCTAAGGCCCTTTAACTCAGATGTCAGGCCACC 2227  
Db 2025 AATGACACCCCAAGTAATCTAGAACATTTCTAAGGCCCTTTAACTCAGATGTCAGGCCACC 2084  
Qy 2228 GGGCAAAACCCGCTCAATACCTCCCAACCAAGGAATGAGATATGTGACCTCACTGCTCCC 2287  
Db 2085 GGGCAAAACCCGCTCAATACCTCCCAACCAAGGAATGAGATATGTGACCTCACTGCTCCC 2144  
Qy 2288 CAACCCAGCGTACGCTGGGACACGCCAAACGCTGTCGGGTTGGAAACAGCAGAGGCTCA 2347  
Db 2145 CAACCCAGCGTACGCTGGGACACGCCAAACGCTGTCGGGTTGGAAACAGCAGAGGCTCA 2204  
Qy 2348 GAAACTGGCTCTGAAATAGGAGACCTTAGCAAGGAGGAGATACAGGGTATCGGGGTTG 2407  
Db 2205 GAAACTGGCTCTGAAATAGGAGACCTTAGCAAGGAGGAGATACAGGGTATCGGGGTTG 2264  
Qy 2408 AGTGTCTCAGAACTCATTTCCGGAAAGATAAATTCAGTGGCTGGCGCAGCACCTGCAATT 2467  
Db 2265 AGTGTCTCAGAACTCATTTCCGGAAAGATAAATTCAGTGGCTGGCGCTCAGCCACCTGCAATT 2324  
Qy 2468 CAAGCTTGGACAGCGGCTTCTTGTGGGAGGCAAAATTTCCCTAGGAAAGAGAGACA 2527  
Db 2325 CAAGCTTGGACAGCGGCTTCTTGTGGGAGGCAAAATTTCCCTAGGAAAGAGAGACA 2384













Qy	305	GGAGGAGAGCGCTGTCTCAGCTGGAGAGGGGTCA	TCCCTGTCTTAAGCCAGTGCCTCTTCC	364
Db	241	GGAGGAGAGCGCTGTCTCAGCTGGAGAGGGGT	CATCCCTGCTTCAAGCCAGTGCCTCTTCC	300
Qy	365	CAGCTCCCATGGGGACCCAGAACCCAGCTCCCGATGGA	AAACGTGGAGCTGAAGGAG	424
Db	301	CAGCTCCCATGGGGACCCAGAACCCAGCTCCCGATGGA	AAACGTGGAGCTGAAGGAG	360
Qy	425	AATGGCAGGACGAGAGATCTTCCAGGCGCACTCC	CAGAAGAGACGGGGGTGGAACTGCTTGG	484
Db	361	AATGGCAGGACTAAGATCTTCCAGGCGCACTCC	CAGAAGAGACGGGGGTGGAACTGCTTGG	420
Qy	485	GCAGCCCGGTGGAGAGACATCTCTCTCCCA	CACGCTAAATTTCAACGGAGCGCGATC	544
Db	421	GCAGCCCGGTGGAGAGACATCTCTCTAACTT	CCAAACGCTAAATTTTCAACGGAGCGCGATC	480
Qy	545	GTAAAGAGGAGACGCTGTGTGGCCCCCAGAGAT	CAACATTTCTCTGGATCAGAGTGAGGGGT	604
Db	481	GCAAGAGGAGACGCTGTGTGGCCCCCAGAGAT	CAACATTTCTCTGGATCAGAGTGAGGGGT	540
Qy	605	CCCTGCTGTCCGATGACTTCTTTGGATACCC	CTGATGACCTGGGATTTAAACGTGGATGACA	664
Db	541	CCCTGCTGTTCGATGACTTCTTTGGATACCC	CTTGGATGACCTGGGATTTAAACGTGGATGACA	600
Qy	665	TCGAGACCCCCGATGAGACCGACTCGCTGGAG	TTCTCTGGGGGAAATGGCAACGAACCTGGAGT	724
Db	601	TCGAGACCCCCGATGAGACCGACTCGCTGGAG	TTCTCTGGGGGAAATGGCAACGAACCTGGAGT	660
Qy	725	GGGAGACGACACCCCGGTGGCCACCGGCCA	AGAAATCATGCCGG	767
Db	661	GGGAGACGACACCCCGGTGGCCACCGGCCA	AGAAATCATGCCGG	703

## RESULT 9

```

US-09-918-995-2540
; Sequence 2540, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2540
; LENGTH: 485
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(485)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-2540

```

Db	184	ACGGGGGTGGAACTGCTTTGGGAGCCCGGTGGGAAGACACATCTCTCTCTCTCCCAACACGCTA	243
Qy	526	AAITTTCAACGGAGCGCATCGTAAGAGAGGAAGACGCTGGTGGCCCCCAGAGATCAACATTTC	585
Db	244	AAITTTCAACGGAGCGCATCGTAAGAGAGGAAGACGCTGGTGGCCCCCAGAGATCAACATTTC	303
Qy	586	CTGGATCAGAGTGAAGGGGTCCTGCTGTCCTCCGATGACTTCTTGGATACCCCTGATGACCTG	645
Db	304	CTGGATCAGAGTGAAGGGGTCCTGCTGTCCTCCGATGACTTCTTGGATACCCCTGATGACCTG	363
Qy	646	GATATTAACTGGATGACATCGAGACACCCCGATGAGACCGACTCGCTGGAGTTTCCTGGGG	705
Db	364	GATATTAACTGGATGACATCGAGACACCCCGATGAGACCGACTCGCTGGAGTTTCCTGGGG	423
Qy	706	AATGGCAACGAATCTGGAGTGGGAAGACGACACCCCCGTGGCCACCGCCAAAGAAATGCCCC	765
Db	424	AATGGCAACGAATCTGGAGTGGGAAGACGACACCCCCGTGGCCACCGCCAAAGAAATGCCCC	483
Qy	766	GG GG 767	
Db	484	GG GG 485	

RESULT 10  
 US-10-723-860-4072  
 ; Sequence 4072, Application US/10723860  
 ; Publication No. US20040253606A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Aziz, Natasha  
 ; APPLICANT: Ginsburg, Wendy M.  
 ; APPLICANT: Zlotnik, Albert  
 ; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &  
 ; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators  
 ; FILE REFERENCE: US5882.0193.NPUS01  
 ; CURRENT APPLICATION NUMBER: US/10/723,860  
 ; CURRENT FILING DATE: 2003-11-26  
 ; PRIOR APPLICATION NUMBER: 60/429,739  
 ; PRIOR FILING DATE: 2002-11-26  
 ; NUMBER OF SEQ ID NOS: 8393  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 4072  
 ; LENGTH: 5656  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-723-860-4072

## RESULT 10

US-10-723-860-4072  
; Sequence 4072, Application US/10723860  
; Publication No. US20040253606A1  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Ginsburg, Wendy M.  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &  
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators  
; FILE REFERENCE: 05882.0193.NPUS01  
; CURRENT APPLICATION NUMBER: US/10/723,860  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: 60/429,739  
; PRIOR FILING DATE: 2002-11-26  
; NUMBER OF SEQ ID NOS: 8393  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4072  
; LENGTH: 5656  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-723-860-4072

QY 829 CTGTGGCGGACAGTATCATCGGGAGCAAGAGCACCGTATAGACCTGCGCATCATGTCGG 888  
Db 1837 CTTTGGAGGACAGTGGTCAATGGAGAACAGAGCAGCGCATTCACATGAAGGTCAATCGAG 1896  
QY 889 CCTTACATGAAGTGGTCAACCGAGGAGTACTACGGCAAGGCTCAAGCGCATCATC 948  
Db 1897 CCTTACAGGAGTCAATTCACGGAGGATCTATGGGACCGTCTAAATGCGCATCAT 1956  
QY 949 GTCTTCGACGCTGCTTCTCCAGACAGCAGCGCTCCCGACTACCACTACATCATGGAG 1008  
Db 1957 GTGTTGCGCGCTGTTCTGCGCAGACAGCAGTGGGCGGATACCACTATGTCATGGAA 2016  
QY 1009 AACCTCTTCTGATCATCAGAGCTTGAAGCTCTGTGTGCTGAGGACTACATGATC 1068  
Db 2017 AATCTTTTCTATATGTAATAGTACTTTAGAGTTGATGCTGAAGACTATATGAT 2076  
QY 1069 GTGTACTGAAGCTGCGCCCGCGGAGGATGCTGGAATCGGCTGGCTGAAGAAG 1128  
Db 2077 GTGTACTGAATGCTGCAACCCCAAGAGGAGATGCGAGGGCTGAGGATGAAGAA 2136  
QY 1129 TGCTACAGATGATCGACCGAGGTGCGGAAAACTGAAGTCTTGTATCATCGTCCAC 1188  
Db 2137 TGCTACAGATGATGACAGCGTGTGAGGAGAGATTTGAATCATCATGTTTCAT 2196  
QY 1189 CCTCTGTGTTCAATCGGACTGTGTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAAGTTC 1248  
Db 2197 CCATCTTGGTTTCAATCAGAACATCTCTGCTGACAGCGCTTTTATAGTCCAAATTC 2256  
QY 1249 ATCAACAGATTCAGTACGTGACAGCTGGAAAGACTGGAGCAACTCATCCCTATGGAA 1308  
Db 2257 AGCAGTAAATTTAAATATGTCAATAGCTTATCAGAACTCAGTGGGCTGATCCCAATGGAT 2316  
QY 1309 CAGTCCAGATCCAGACTGCTGCTGCAATACGAGAGGAA 1350  
Db 2317 TGCATCCACATTCAGAGAGCATCATCAAACTGGATGAAGAA 2358

## RESULT 11

US-10-756-149-4047  
; Sequence 4047, Application US/10756149  
; Publication No. US20050181375A1

; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha

; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND

; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER  
; FILE REFERENCE: file  
; CURRENT APPLICATION NUMBER: US/10/756,149

; CURRENT FILING DATE: 2004-01-12  
; NUMBER OF SEQ ID NOS: 5818

; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4047

; LENGTH: 5656  
; TYPE: DNA

; ORGANISM: Homo Sapiens  
US-10-756-149-4047

Query Match  
Best Local Similarity 63.3%;

Matches 558; Conservative 0; Mismatches 291; Indels 33; Gaps 2;  
QY 469 GGGGTGGAAGTCTTGGAGCGCGGTGGAGACACATCTCTCTCCCGACAGCGTAAT 528

Db 1510 GAGGGGTGCTGAGTCCCGAGTCTGAGACATGAGCGCTGGAACCTCTTAATCTCTGGAT 1569

QY 529 TTCAACGGAGCGCATCGTAAAGAGAGACGCTGGTGGCGCCAGAGATCAACATTTCTCTG 588

Db 1570 CTTAATGACACTCATCTCGAGATCAAGCTCAGAGCCCAATATCATCTTCTCTG 1629

QY 589 GATCAGATGAGGGGTCTCTGTCGCGATGACTTTCTGGATACCCCTGATGACCTGGAT 648

Db 1630 GACCAAGTGAAGGATCTATTCTCTGATGATACTTGGACAGTCCAGATGAAATTCAG 1689

QY 649 ATTAACTGTGATGACATCGAGACCCCGGATGAGACCGACTCGCTGGAGTTCTTGGGAAAT 708  
Db 1690 ATCAATGTGTGATGAACTTGTATCCCCCGGATGAGCAGATTTCTTTGAGTACTACTGGCCAT 1749  
QY 709 GGCACGAACACTGGAGTGGGAAGACGACACCCCGGTGGCCACCGCCCAAGAACATGTCGCCGG 768  
Db 1750 G-----AAGATCCACAGCCCAACAAAGATTCTGGC 1779  
QY 769 GACAGCGCGATCTATTTGGGACCGCAGACGGAGGACGGCAGCGCCGCCCAACCGGGCGC 828  
Db 1780 CAAGAGTCAAGTCTATTTCCAGAAATATACGGCCGAGAGAGAAACGGGAGGACAAAC---CG 1836  
QY 829 CTGTGCGGACAGTCAATCATCGGGAGCAAGAGCAGCTGATAGACTGCTGACATGATGATCCCG 888  
Db 1837 CTTTGGAGGACAGTGGTCAATGGAGAACAGAGCAGCGCATTTGACATGAAGGTCAATCGAG 1896  
QY 889 CTTTACATGAAGTGGTCAACCGAGGATCTACGGGAAAGGCTCAACGCGCATCATC 948  
Db 1897 CCTACAGGAGTCAATTTCTCAGCGAGGATCTATGGGACGGTCTAATGCCATCAT 1956  
QY 949 GTCTTGGCAGCGCTTCTTCCAGACAGCAGCGCTCCCGCATACCACTACATCATGGAG 1008  
Db 1957 GTGTTTGGCGCGCTGTTTCTGCGCAGACAGCAGTCCGGCGGATTTACCACTATGTCATGGAA 2016  
QY 1009 AACCTCTTCTGATCATCAGCAGCTTAGAGCTCTGCTGGCTGAGGACTACATGATC 1068  
Db 2017 AATCTTTTCTATATGTAATAGTACTTTAGAGTTGATGCTGTAAGACTATATGAT 2076  
QY 1069 GTGTACTGAAAGTGGCCACCGCGGGAGGATGCTTGGAAATCGGCTGGCTGAAGAAG 1128  
Db 2077 GTGTACTTGAATGGTGCACCCCAAGAGAGGAGATGCGAGGCTGAGGCTGGATGAAGAA 2136  
QY 1129 TGCTACAGATGATCGACCGAGGTGCGGAAAACTTGAAGTCTCTTGTATCATCGTCCAC 1188  
Db 2137 TGCTACAGATGATGACAGCGTGTGAGGAGATTTGAATCATCATGTTTCAT 2196  
QY 1189 CCTCTGTGTTCAATCGGACTGTGTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAAGTTC 1248  
Db 2197 CCATCTTGGTTTCAATCAGAACATCTCTGCTGACAGCGCTTTTATAGTCCAAATTC 2256  
QY 1249 ATCAACAGATTCAGTACGTGACAGCTGGAAAGACTGGAGCAACTCATCCCTATGGAA 1308  
Db 2257 AGCAGTAAATTTAAATATGTCAATAGCTTATCAGAACTCAGTGGGCTGATCCCAATGGAT 2316  
QY 1309 CAGTCCAGATCCAGACTGCTGCTGCAATACGAGAGGAA 1350  
Db 2317 TGCATCCACATTCAGAGAGCATCATCAAACTGGATGAAGAA 2358

## RESULT 12

US-10-487-593-2

; Sequence 2, Application US/10487593

; Publication No. US20050064413A1

; GENERAL INFORMATION:

; APPLICANT: NAKAGAWARA, AKIRA

; APPLICANT: HATTORI, MASAHIRA

; APPLICANT: SAKAKI, YOSHIYUKI

; TITLE OF INVENTION: NOVEL HUMAN BMCC1 GENE

; FILE REFERENCE: 7388/80977

; CURRENT APPLICATION NUMBER: US/10/487,593

; CURRENT FILING DATE: 2004-02-24

; PRIOR APPLICATION NUMBER: PCT/JP02/08520

; PRIOR FILING DATE: 2002-08-23

; PRIOR APPLICATION NUMBER: JP 2001-255198

; NUMBER OF SEQ ID NOS: 55

; SOFTWARE: PatentIn ver. 3.2

; SEQ ID NO 2

; LENGTH: 11690

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-487-593-2

```
Query Match      11.2%; Score 327.6; DB 9; Length 11690;
Best Local Similarity 63.2%; Pred. No. 2.2e-83;
Matches 557; Conservative 0; Mismatches 289; Indels 36; Gaps 2;

QY 469 GGGGTGGAAGTCTTGGCAGCCGGTGGAGACACATCTCTCTCCCAACAGCTAAAT 528
DB 7547 GAGGGCGTGTGAGTCCAGTGTGCAGACATGAGCGCTGAACTCTTAATCTCTGGAT 7606
QY 529 TTCAAGCGAGCGCATCTTAAGAGAGAGACGCTGGTGGCCCCAGAGATCAACATTTCTCTG 588
DB 7607 CTTAATGACACTCATCTCTCGAGAACTCAGCTCAGAGCCCCAATATCAATCTTTCTCTG 7666
QY 589 GATCAGAGTGAAGGGTCCCTGTGTGCGATGACTTTTGTGATACCCCTGATGACCTGGAT 648
DB 7667 GACCAAGTGAAGGATCTATTCTCTGTGATGATACTTGGACAGTCCAGATGAATGAC 7726
QY 649 ATTAAGTGTGATCAGATCGAGACCCCGATGAGACCGACTCGTGGAGTCTCTGGGAAT 708
DB 7727 ATCAATGTGATGAACCTTGATACCCCGATGAAGCAGATTTCTTTGAGTACACTGG- 7782
QY 709 GGCAACGAAGTGGAGGAGACGACACCCCGTGGCCACCGCCGCAAGAACATGCCCGG 768
DB 7783 -----CCATGATCCACAGCAAAAGATTTCTGGC 7813
QY 769 GACAGCGCGGATCTATTTTGGGAGCGGCACGAGAGGAGCGGAGCGCCGCCAACGGGCGC 828
DB 7814 CAAGAGTCAAGTCTATTCCAGAAATATAGCGCGAAGAGGAAACGGGAGGACAAAC---CG 7870
QY 829 CTGTGGCGGACATGATCATCGGGGAGCAAGACACGATATAGACCTGACATGATCCGG 888
DB 7871 CTTTGGAGGACAGTGGTCAATTTGAGAAACAGAGCAGCGCAATGACATGAAGTCACTGAG 7930
QY 889 CTTTACATGAAGTGGTCAACCGAGGAGTACTACGGGAGGAGCGCTCAACGCCATCATC 948
DB 7931 CCTTACAGGAGATCAATTTCTACGAGGAGATCTATGGGACGGTCTAAATGCCATCAT 7990
QY 949 GTCTTCCGAGCTGTCTTCTTCCAGACAGCAGCGCTCCCGACTACCACTACATCATGGAG 1008
DB 7991 GTGTTTGGCCGCTGTTTCTGCCAGACAGCAGTCCGGCGGATTTACCACATGATCATGAA 8050
QY 1009 AACCTTCTCTGTACGTATCAGCAGCTTATAGACTCTGTGTGCTGAGACTATCATGATC 1068
DB 8051 AATCTTTTCTATATGTAATGAATCACTTTAGAGTTGATGTAGTCTGAAGACTATATGAT 8110
QY 1069 GTGTACTCTGAACCGTGCACCGCCCGGAGGATGCTTGGAAATCGGCTGGTGAAGAAG 1128
DB 8111 GTGTACTTGAATGGTGCACCCCAAGAGGAGATGCCAGGGCTAGGCTGGATGAAGAA 8170
QY 1129 TGCTACAGATGATCGACCGAGGTTTGGGAAACCTTGAAGTCTTGTGATCATCGTCCAC 1188
DB 8171 TGCTACAGATGATGACAGACGTTTGGAGAAATTTGAAATCATTCATCTGTTTCAAT 8230
QY 1189 CCCTCGTGGTTCAATTCGAGTGTGTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTC 1248
DB 8231 CCATCTTGGTTTCAACGAACATCTTGTGTGACACGACCTTTTATAAGTTTCAAAATTC 8290
QY 1249 ATCAACGAAGTCCAGTACGTGACAGCTTGGAGAGCTGGAGCACTCATCTCCATATGAA 1308
DB 8291 AGCAGTAAATTAATATATGATCAATAGCTTATCAGAACTCAGTGGGCTGATCCCAATGGAT 8350
QY 1309 CAGCTCCAGATCCAGATCGCTCTGCAATATCGAAGAGAA 1350
DB 8351 TGCATCCACATTCAGAGACATCATCAAACTGATGAAGAA 8392
```

## RESULT 13

```
US-09-873-367C-644
; Sequence 644, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
```

```
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 644
; LENGTH: 5654
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-873-367C-644
```

```
Query Match      10.9%; Score 318.2; DB 3; Length 5654;
Best Local Similarity 63.2%; Pred. No. 8.7e-81;
Matches 558; Conservative 0; Mismatches 288; Indels 37; Gaps 3;

QY 469 GGGGTGGAAGTCTTGGCAGCCGGTGGAGACACATCTCTCTCTCCCAACAGCTAAAT 528
DB 1510 GAGGGCGTGTGAGTCCAGTCCAGTGTGCAGACATGAGCGCTGAACTCTTAATCTCTGGAT 1569
QY 529 TTCAAGCGAGCGCATCTTAAGAGGAGAGCGTGGTGGCCCCAGAGATCAACATTTCTCTG 588
DB 1570 CTTAATGACACTCATCTCTCGGAGAAATCAAGCTTCAAGCCCCCAATATCAATCTTTCTCTG 1629
QY 589 GATCAGAGTGAAGGGTCTCTGTCTCCGATGACTTTCTTGGATACCCCTGATGACCTGGAT 648
DB 1630 GACCAAGTGAAGGATCTATTCTCTGATGATACTTTGGACAGTCCAGATGAATTGAC 1689
QY 649 ATTAAGTGTGATGATCATCGAGACCCCGATGAGACCGACTCGCTGGAGTCTCTGGGGAAT 708
DB 1690 ATCAATGTGATGAATCTTGATACCCCGATGAAGCAGATTTCTTTGAGTACACTGG- 1745
QY 709 GGCAACGAAGTCTGAGTGGGAGACGACACCCCGTGGCCACCGCCCAAGAACATGCCCGG 768
DB 1746 -----CCATGATCCCAAGCCCAACAAAGATTTCTGGC 1776
QY 769 GACAGCGGGATCTATTTTGGGAGCGGCACGAGAGGAGCGGAGCGCCGCCAACGGGCGC 828
DB 1777 CAAGAGTCAAGTCTATTTCAGAAATATACGGCCGGAAGAGGAAACGGGAGGACAAAC---CG 1833
QY 829 CTGTGGGAGACAGTATCATCTGGGAGCAAGAGCACCGTATAGACTGACATGATCCGG 888
DB 1834 CTTTGGAGGACAGTGGTCAATTCAGGAAACAGAGACGCGCATTGAATGAAGGTCTATCGAG 1893
QY 889 CCTTACATGAAGTGGTCAACCCAGCGGGTACTACGGGAGGCGCTCAACGCCATCATC 948
DB 1894 CCTTACAGAGAGTCAATTTCTCAGGAGGATATATGGGAGCGGCTCTAAATGCCATCAT 1953
QY 949 GTCTTCCGAGCTGTCTTCTTCCAGACAGCAGCTCCCGACTACCACTATCATATGAGAG 1008
DB 1954 GTGTTTGGCGCTGTTTCTGCCAGACAGCAGTCCGGGCGGATTAACATATGATGAA 2013
QY 1009 AACCTTCTCTGTATCGTATCAGAGCTTATAGCTCTCTGTGGTGGCTGAGAGCTATCATGATC 1068
DB 2014 AATCTTTTCTATATGTAATGAATCTTTAGAGTTGATGGTACGTGAAGACTATATGAT 2073
QY 1069 GTGTACCTGMAAGGTCACCGCC- CGGCGGAGGATGCTCGGAATCGGCTGGCTCAAGAA 1127
DB 2074 GTGTACTTGAATGGTGCACACCCCAAGAGGAGGATGCCAGGGCTAGGCTGGATGAAGAA 2133
```

QY 1128 GTCTACCAAGATGATCGACCGGAGTTGCGGAAACCTGAACTCTTGATCATCTGTCCTCA 1187  
DB 2134 ATGCTACCAAGATGATGACAGCGTTGAGGAAAGAAATTTGAAATCATTTATCTTCA 2193  
QY 1188 CCCCTCGTGTTCATTCGAGCTGTGCTGGCCATCTCTCGCCCTTTTCATCAGCTCAAGTT 1247  
DB 2194 TCCATCTTGTTCATCAGAACATCTTCTGTGTCAGACGACCTTTTATAGTTCAAAAT 2253  
QY 1248 CATCAACAGATCCAGTACGTCGACAGCTTTGGAAGACCTGGAGCAACTCATCTCTATGGA 1307  
DB 2254 CAGCAGTAAATTAATATGCTCAATAGCTTATCAGAACTCAGTGGGCTGATCCCAATGGA 2313  
QY 1308 ACAGTCCAGATCCAGATGCTGCTCTGCAATACGAAGAGGAA 1350  
DB 2314 TTGCATCCACATTTCCAGAGAGCATCATCAAACTGGATGAAGAA 2356

## RESULT 14

US-09-968-007A-349  
; Sequence 349, Application US/09968007A  
; Publication No. US20040115625A1  
; GENERAL INFORMATION:  
; APPLICANT: Ebner, Reinhard  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signa  
; FILE REFERENCE: 689290-71  
; CURRENT APPLICATION NUMBER: US/09/968, 007A  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 1001  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 349  
; LENGTH: 5654  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-968-007A-349

Query Match 10.9%; Score 318.2; DB 3; Length 5654;  
Best Local Similarity 63.2%; Pred. No. 8.7e-81;  
Matches 558; Conservative 0; Mismatches 288; Indels 37; Gaps 3;  
QY 469 GGGGTGGAATGCTTGGCAGCCGGTGGAGACATCTCTCTCCCAACAGCTAAAT 528  
DB 1510 GAGGGCGTCTGAGTCCCAAGTGTGAGACATGAGGCGCTCTCTAAATCTCTGGAT 1569  
QY 529 TTCAACGAGCGCATCTGTAAGAGAGACGCTGTGTCGCCCCAGAGATCAACATTTCTCTG 588  
DB 1570 CTTAATGACACTCATCTCTCGAGATCAAGCTCAGAGCCCAATATCAATCTTCTCTG 1629  
QY 589 GATCAGAGTGAGGGTCCCTGCTGTCGATGATCTCTTGGATACCCCTGATGACCTGGAT 648  
DB 1630 GACCAAGTGAAGGATCTATCTCTCTGATGATACTTGGACAGTCCAGATGAATGAC 1689  
QY 649 ATTAAGCTGATGATCAGATCCAGACCCCGGATGAGACCGACTCGTGGAGTTCCTGGGAAAT 708  
DB 1690 ATCAATGTGGATGAATCTGATATCCCGGATGAGCAGATCTTTTGGATACACTGG- 1745  
QY 709 GGCAACGAATGAGTGGGAGAGACACCCCGTGGCCACCGCCAGAACATGCCCGG 768  
DB 1746 -----CCATGATCCACAGCCCAAGAAATTTCTGTC 1776  
QY 769 GACAGCGCGATCTATTTGGGAGCGCAGCAGGAGGCGGCGCGCCCAACGCGCGC 828

DB 1777 CAAGAGTCAGAGCTATTTCCAGAAATATACGGCCGAGAGAAACGGGAGCAAC- 1833  
QY 829 CTGTGCGGACAGTATCATTCGGGAGCAGAGCAGCCGTATAGACTGACATGATCCGG 888  
DB 1834 CTTTGGAGACAGTGGTCAATGGAGAACAGAGCAGCGCATTTGACATGAAGGTCTATCGAG 1893  
QY 889 CCTTACATCAAAAGTGGTCAACCCAGGAGGTACTACGGGAAAGGCTCAACGCCATCATC 948  
DB 1894 CCTTACAGAGAGTCAATTTCTCAGGAGGATATATGGGAGCGGTCTAATGCCATCAT 1953  
QY 949 GTCTTGGAGCTGCTTCTTCCAGACAGACCTCCCGACTACCACTACATCATCATGAG 1008  
DB 1954 GTGTTTGGCGCTGTTTCTGCGCAGACAGCAGTCCGGGGATTACCACTATGTCATGGA 2013  
QY 1009 AACCTCTTCTGTCATGTCATCAGCAGCTTAGAGCTCTGCTGGCTGAGGACTACATGATC 1068  
DB 2014 AATCTTTTCTATATGTAATAAGTACTTTAGAGTTGATGGTAGCTGAAGACTATATGAT 2073  
QY 1069 GTGTACCTGAACGGTGCACAGCC- CCGCGGAGGATGCTGGAATCGGCTGCTGAAGAA 1127  
DB 2074 GTGTACTTGAATGGTGCACACCCCAAGAGAGGAGATGCCAGGCTAGGCTGGATGAAGAA 2133  
QY 1128 GTGTACCAAGATGATCGACCGGAGGTTCGGGAAACCTGAAAGTCTTGTATCATCTGTC 1187  
DB 2134 ATGCTACCAAGATGATTCAGACAGCGTTGAGGAAAGAAATTTGAAATCATTCATCTTCA 2193  
QY 1188 CCCCTCGTGTTCATTCGAGCTGTGCTGGCCATCTCTGCGCCCTTTTCATCAGGCTCAAGTT 1247  
DB 2194 TCCATCTTGGTTTCATCAGAACAAATCTTCTGTGACAGACCTTTTATAAGTTCAAAAT 2253  
QY 1248 CATCAACAGATCCAGTACGTCGACAGCTTGGAGACCTTGGAGCAACTCATCTCTATGGA 1307  
DB 2254 CAGCAGTAAATTAATATGTCATAGCTTATCAGAACTCAGTGGGCTGATCCCAATGGA 2313  
QY 1308 ACAGTCCAGATCCAGATGCTGCTCTGCAATACGAAGAGGAA 1350  
DB 2314 TTGCATCCACATTTCCAGAGAGCATCATCAAACTGGATGAAGAA 2356

## RESULT 15

US-09-968-007A-729  
; Sequence 729, Application US/09968007A  
; Publication No. US20040115625A1  
; GENERAL INFORMATION:  
; APPLICANT: Ebner, Reinhard  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sig  
; FILE REFERENCE: 689290-71  
; CURRENT APPLICATION NUMBER: US/09/968, 007A  
; CURRENT FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US/60/237, 172  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: US/60/237, 173  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: US/60/237, 278  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: US/60/237, 294  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: US/60/237, 295  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: US/60/237, 316  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 1001  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO: 729  
; LENGTH: 5654  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-968-007A-729

Query Match 10.9%; Score 318.2; DB 3; Length 5654;  
Best Local Similarity 63.2%; Pred. No. 8.7e-81;

```
Matches 558; Conservative 0; Mismatches 288; Indels 37; Gaps 3;
QY 469 GGGGTGGAACTGCTTGGCAGCCCGTGGAGACACATCTCTCTCCCAACACGCTAAAT 528
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1510 GAGGGCGTGTGAGTCCAGTGTGCGAGACATGAGGCTTGAACCTCTTAATCTCTGGAT 1569
QY 529 TTCAACGGAGCCGATCGTAAGAGGAAGACGCTGGTGGCCCGAGAGATCAACATTTCTCTG 588
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1570 CTTAATGACACTCATCTCGAGAAATCAAGCTCACAGCCCCCAATATCAATCTTTCTCTG 1629
QY 589 GATCAGAGTGAGGGTCCCTGCTGCTCGATGACTTCTTGGATACCCCTGATGACCTGGAT 648
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1630 GACCAAGTGAAGGATCTATTTCTCTGATGATACTTGGACAGTCCAGATGAAATGAC 1689
QY 649 ATTAACGTGATGACATCGAGACCCCGATGAGCCGACTCGCTGGAGTTCTCTGGGAAT 708
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1690 ATCAATGTGGATGACTTGTATACCCCGATGAGCAGATTTCTTTGAGTACACTGG---- 1745
QY 709 GGCAACGAACGTGGAGTGGGAAGACACACCCCGTGGCCACCGCCCAAGAACATGCCCGGG 768
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1746 -----CCATGATCCACAGCCCAAGAAATCTGGC 1776
QY 769 GACAGCGCGGATCTATTTGGGACGGCAGACGGAGGACGGCAGCGCCGCCAACGGGCGC 828
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1777 CAAGGTGAGGTCTATTCAGAAATATACGGCCGAAGAGAACGGGAGGACAAAC---CGG 1833
QY 829 CTGTGGCGGACAGTGTATCTGGGGAGCAAGAGCACCGCTATAGACCTGCACATGATCCGG 888
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1834 CTTTGGAGGACAGTGTCTTGGAGAACAGAGCAGCGGCAATGACATGAGGTATCGAG 1893
QY 889 CTTTACATGAAGTGTGACCCACCGAGGGTACTACGGGGAAGGCTCAACGCCATCATC 948
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1894 CCCTACAGGAGATCATTTCTCACGAGGATATATATGGGACGGTCTAAATGCCATCAT 1953
QY 949 GTCTTCGACGCTGCTTCTTCCAGACAGCAGCGCTCCCGACTTACCACTACATCATGGAG 1008
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1954 GTGTTTGGCGCCTGTTTCTGCGCAGACAGCAGTCCGGCGGATTTACCACATATGTCATGGA 2013
QY 1009 AACCTCTTCTGTACGTCTATCAGCAGCTTAGAGCTCTCTGGTGGCTGAGACTACATGATC 1068
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2014 AATCTTTTCTATATGTATAGTACTTTTAGAGTTGATGGTAGCTGAGACTATATGATT 2073
QY 1069 GTGTACTGAAACGGTGCAGCGCC- CGGCGGAGATGCTTGGAAATCGGCTGGCTGAAGAA 1127
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2074 GTGTACTTGAATGGTGCACCCCAAGAGGAGATGCCAGGGCTAGGCTGGATGAAGAA 2133
QY 1128 GTGCTACCAAGATCGACCGGAGGTTGGGAAAACCTGAACTGAGTCTTGATCATCGTCCA 1187
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2134 ATGCTACCAAGATTTGACAGACGGTTGAGGAAGAAATTTGAAATCATTCATTTGTTCA 2193
QY 1188 CCCTCGTGTTCATTCGGACTGTGCTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTT 1247
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2194 TCCATCTTGGTTCATCAGAACATCTTGTGTCACGACCTTTTATAGTTCAAAAT 2253
QY 1248 CATCAACAGATCCAGTACGTGACAGCTTGGAGACCTGGAGCAACTCATCCCTATGGA 1307
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2254 CAGCAGTAAATTAATATGTCAATAGCTTATCAGAACTCAGTGGGCTGATCCCAATGGA 2313
QY 1308 ACAGTCCAGATCCAGAGCTGGTCTGCAATACGAAGGAA 1350
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2314 TTGCATCCACATTCACAGAGCATCATCAAACTGGATGAAGAA 2356
```

Search completed: March 5, 2006, 05:22:00  
Job time : 2067 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 03:46:55 ; Search time 810 Seconds  
(without alignments)

7898.360 Million cell updates/sec

Title: US-10-699-941-3

Perfect score: 2918

Sequence: 1 gcggagctctgcgcgcct.....tctcaaaaaaaaaaaaaa 2918

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7218535 seqs, 1096242582 residues

Total number of hits satisfying chosen parameters: 14437070

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US08 NEW PUB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/US06 NEW PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US07 NEW PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/PCT NEW PUB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US09 NEW PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US09 NEW PUB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US10 NEW PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US10 NEW PUB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US11 NEW PUB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US11 NEW PUB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US11 NEW PUB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US11 NEW PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60 NEW PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2598.2	89.0	2759	9	US-11-072-512-997
2	1325.8	45.4	2625	9	US-11-072-512-986
C 3	243.6	8.3	623	6	US-09-925-065A-515679
C 4	241.6	8.3	403278	8	US-10-995-561-13421
5	241.4	8.3	560	6	US-09-925-065A-594866
C 6	240.8	8.3	138827	7	US-10-330-773-159
7	239.8	8.2	560	6	US-09-925-065A-594865
8	239.6	8.2	628	6	US-09-925-065A-55092
9	239.2	8.2	611	6	US-09-925-065A-482665
C 10	239.2	8.2	611	6	US-09-925-065A-563212
C 11	239.2	8.2	67126	8	US-10-995-561-13342
C 12	239.2	8.2	127340	12	US-11-112-908-35
C 13	239	8.2	483	6	US-09-925-065A-621024
C 14	239	8.2	1082	6	US-09-925-065A-84578
C 15	238.6	8.2	2649	6	US-09-925-065A-723891
C 16	238.4	8.2	547	6	US-09-925-065A-657480
C 17	238.4	8.2	623	6	US-09-925-065A-687639
C 18	238.4	8.2	623	6	US-09-925-065A-687640
C 19	238.2	8.2	624	6	US-09-925-065A-849333
C 20	238	8.2	180531	9	US-11-114-798-57

C 21	237.8	8.1	694	6	US-09-925-065A-75055
C 22	237.8	8.1	1288	6	US-09-925-065A-692997
C 23	237.6	8.1	611	6	US-09-925-065A-563211
24	237.4	8.1	482	6	US-09-925-065A-624538
25	237.4	8.1	482	6	US-09-925-065A-624539
26	237.4	8.1	570	6	US-09-925-065A-197384
27	237.2	8.1	621	6	US-09-925-065A-337721
28	236.6	8.1	883	6	US-09-925-065A-8768
29	236.2	8.1	883	6	US-09-925-065A-486749
30	236.2	8.1	883	6	US-09-925-065A-8769
C 31	236.2	8.1	187745	12	US-11-121-086-83
C 32	236	8.1	600	6	US-09-925-065A-873038
C 33	236	8.1	661	6	US-09-925-065A-94635
C 34	236	8.1	1980	6	US-09-925-065A-92083
C 35	236	8.1	1980	6	US-09-925-065A-596162
C 36	236	8.1	171486	12	US-11-121-086-105
C 37	236	8.1	179777	12	US-11-121-086-106
C 38	235.8	8.1	199321	12	US-11-121-086-10
C 39	235.6	8.1	661	6	US-09-925-065A-94636
C 40	235.6	8.1	3125	9	US-11-072-512-1355
C 41	235.6	8.1	40000	8	US-10-995-561-13512
C 42	235.4	8.1	196200	12	US-11-121-086-9
C 43	235.2	8.1	574	6	US-09-925-065A-443115
44	235	8.1	553	6	US-09-925-065A-930747
45	235	8.1	580	6	US-09-925-065A-282862

#### ALIGNMENTS

##### RESULT 1

US-11-072-512-997  
; Sequence 997, Application US/11072512  
; Publication No. US20060029945A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHICO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOTOKU  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: Novel full length cDNA  
; FILE REFERENCE: 084335-0191  
; CURRENT APPLICATION NUMBER: US/11/072,512  
; CURRENT FILING DATE: 2005-03-07  
; PRIOR APPLICATION NUMBER: US 60/350,978  
; PRIOR FILING DATE: 2002-01-25  
; PRIOR APPLICATION NUMBER: JP 2001-379298  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 997  
; LENGTH: 2759  
; TYPE: DNA  
; ORGANISM: Homo sapiens

Query Match 89.0%; Score 2598.2; DB 9; Length 2759;  
Best Local Similarity 95.2%; Pred. No. 7.5e-25;  
Matches 2756; Conservative 0; Mismatches 3; Indels 136; Gaps 2;

QY 8 CTCTGCCAGCCCTGAGCTGGGAAGAACAGCTACCTCGGAGGCGGCGCGAGCGGCG 67  
DB 1 CTCTGCCAGCCCTGAGCTGGGAAGAACAGCTACCTCGGAGGCGGCGCGAGCGGCG 60  
QY 68 GCGGATGAGAGGGGGCGAGCGCGAGCCCGGCTGGGAGGCCCAACCGCTAACCCCTGCAC 127  
DB 61 GCGGATGAGAGGGGGCGAGCGCGAGCCCGGCTGGGAGGCCCAACCGCTAACCCCTGCAC 120  
QY 128 CCACCCACCCCTGCAACAAAGAGCTGGGGGCGCTGGGCCACGTCGCCCTGGGTGACCTT 187  
DB 121 CCCACCCA----- 128  
QY 188 CCTCGGATGAGAAATCCGCGCCCTGCGAGCAATCTCTTCTCTAGGCTCTGAAGGCCCGG 247  
DB 129 -----CCCTGCGAGCATCTCTTCTCTAGGCTCTGAAGGCCCGG 170  
QY 248 GGAGCGTGAGCGATGCCAGCTGCACCCGCGGAGGGCTCGCCCTTGTGTCAGTAAGGA 307  
DB 171 GGAGCGTGAGCGATGCCAGCTGCACCCGCGGAGGGCTCGCCCTTGTGTCAGTAAGGA 230  
QY 308 GGAGGCTGTCTCAGCTGCAGAGGGGTCAATCCCTGCTTCAAGCCAGTGCCTCTTCCCAG 367  
DB 231 GGAGGCTGTCTCAGCTGCAGAGGGGTCAATCCCTGCTTCAAGCCAGTGCCTCTTCCCAG 290  
QY 368 CTCCTATGGGGACCAACGAGCCACGCTCGGATGGAACAGTGGACGTGAAGGAGGAT 427  
DB 291 CTCCTATGGGGACCAACGAGCCACGCTCGGATGGAACAGTGGACGTGAAGGAGGAT 350  
QY 428 GCGAGGACGAGATCTTCCAGGCCACTCCAGAGAGAGCGGGGTGGAACTGCTTGGCA 487  
DB 351 GGC----- 353  
QY 488 GCCCGGTGAAGACACATCTCTCTCCCAACAGCTAAATTTCAACGAGGCGCATCGTA 547  
DB 354 -----AAGACACATCTCTCTCTCCCAACAGCTAAATTTCAACGAGGCGCATCGTA 404  
QY 548 AAGAGAGAGCGTGTGGGCCCAAGAGATCAACATTTCTTGGATCAGAGTGGGGTCCC 607  
DB 405 AAGAGAGAGCGTGTGGGCCCAAGAGATCAACATTTCTTGGATCAGAGTGGGGTCCC 464  
QY 608 TGCCTGTCGATGACTTCTTGGATACCCCTGATGACCTGGATATTAACGTGGATGACATCG 667  
DB 465 TGCCTGTCGATGACTTCTTGGATACCCCTGATGACCTGGATATTAACGTGGATGACATCG 524  
QY 668 AGACCCCGATGAGACCGACTCGCTGGAGTTCTGGGGAAATGGCAACGAACTGGAGTGG 727  
DB 525 AGACCCCGATGAGACCGACTCGCTGGAGTTCTGGGGAAATGGCAACGAACTGGAGTGG 584  
QY 728 AAGACGACACCCCGTGGGCCACCGCCAGAAACATGCCCCGGGGACAGCGGATCTATTGG 787  
DB 585 AAGACGACACCCCGTGGGCCACCGCCAGAAACATGCCCCGGGGACAGCGGATCTATTGG 644  
QY 788 GGGAGCGCACGAGGAGGAGCGAGCGCCGCAACGGGCGCTGTGGCGGACAGTGATCA 847  
DB 645 GGGAGCGCACGAGGAGGAGCGAGCGCCGCAACGGGCGCTGTGGCGGACAGTGATCA 704  
QY 848 TCGGGGAGCAAGACACCGTATAGACTGCAATGATCCGGCTTACATGAAGTGGTCA 907  
DB 705 TCGGGGAGCAAGACACCGTATAGACTGCAATGATCCGGCTTACATGAAGTGGTCA 764  
QY 908 CCCACGAGGATCTACGGGGAAGGCTCAACGCCCATCATCGTCTTCGAGCGCTGCTTCC 967  
DB 765 CCCACGAGGATCTACGGGGAAGGCTCAACGCCCATCATCGTCTTCGAGCGCTGCTTCC 824  
QY 968 TTCCAGACGAGCGCTCCCGGACTACCATCATCATGAGAACCTTCTCTGTAGCTCA 1027  
DB 825 TTCCAGACGAGCGCTCCCGGACTACCATCATCATGAGAACCTTCTCTGTAGCTCA 884  
QY 1028 TCAGCAGCTTAGAGCTCTCTGGTGGCTGAGGACTACATGATCGTGTACCTGAAACGGTGCCA 1087  
DB 885 TCAGCAGCTTAGAGCTCTCTGGTGGCTGAGGACTACATGATCGTGTACCTGAAACGGTGCCA 944  
QY 1088 CGCCCCGCGGAGGATGCTCGGAACTCGGCTGGCTGGAAGAGTGCTTACCAAGATGATCGACC 1147

DB 945 CGCCCCGCGGAGGATGCTCGGAATCGGCTGGCTGAAGAGTGCTACCAAGATGATCGACC 1004  
QY 1148 GGAGCTTGGGNAACCTGGAAGTCTTGTATCATCGTCCACCCCTCGTGGTTTCAATTCGGA 1207  
DB 1005 GGAGTGTGGGNAACCTGGAAGTCTTGTATCATCGTCCACCCCTCGTGGTTTCAATTCGGA 1064  
QY 1208 CTGTCTCGCCCATCTCTCGCCCTTTCATCAGCGTCAAGTTTCATCAACAGATCCAGTAGC 1267  
DB 1065 CTGTCTCGCCCATCTCTCGCCCTTTCATCAGCGTCAAGTTTCATCAACAGATCCAGTAGC 1124  
QY 1268 TGCAAGCTTGGAGACCTGGAGCAATCATCTCTTATGAAACACGTCAGATCCAGACT 1327  
DB 1125 TGCAAGCTTGGAGACCTGGAGCAATCATCTCTTATGAAACACGTCAGATCCAGACT 1184  
QY 1328 GGGTCTCTGCATACCAAGAGGAAAGACTGNAAGGCCAGGAGGAGAGCGGAGGCCCGCAGC 1387  
DB 1185 GGGTCTCTGCATACCAAGAGGAAAGACTGNAAGGCCAGGAGGAGAGCGGAGGCCCGCAGC 1244  
QY 1388 CGGAGTGTGTGCTGCCAGGCTCTGAAAGAGAGCCAGAGGTGGCAACAGTGGAAACACAGGT 1447  
DB 1245 CGGAGTGTGTGCTGCCAGGCTCTGAAAGAGAGCCAGAGGTGGCAACAGTGGAAACACAGGT 1304  
QY 1448 CTGTCTCTGTCTCAGAGATCAGGAAACAAAGCATGTCTTGAGGCGACGTCAGCATAAACAA 1507  
DB 1305 CTGTCTCTGTCTCAGAGATCAGGAAACAAAGCATGTCTTGAGGCGACGTCAGCATAAACAA 1364  
QY 1508 AGGACATGGNAGAGATTCAGATGCCAGAAACCTCTGTGAGCGCCACCTGGGCCCGAG 1567  
DB 1365 AGGACATGGNAGAGATTCAGATGCCAGAAACCTCTGTGAGCGCCACCTGGGCCCGAG 1424  
QY 1568 ATCTCATCTCTGCTCATCTGAGTCCCAATCTTCAAGGGTGGCCAGCCCTCGTTCATC 1627  
DB 1425 ATCTCATCTCTGCTCATCTGAGTCCCAATCTTCAAGGGTGGCCAGCCCTCGTTCATC 1484  
QY 1628 TCTGAAACCCAGCATCTCTTTTTCAGCTGTGTGAAACATTTGTATTTTTTTTTTAAACGAT 1687  
DB 1485 TCTGAAACCCAGCATCTCTTTTTCAGCTGTGTGAAACATTTGTATTTTTTTTTTAAACGAT 1544  
QY 1688 GCGATTTTGTGCGTTCCAGAAAGGGGCCAGCTCTGAGCCCTTCAACCTTCCACATCA 1747  
DB 1545 GCGATTTTGTGCGTTCCAGAAAGGGGCCAGCTCTGAGCCCTTCAACCTTCCACATCA 1604  
QY 1748 CGAACTCTCAGCCGAGGAAAGCAAGCGGAGGGGTGGCCCGGCTGGCTGGTGGCC 1807  
DB 1605 CGAACTCTCAGCCGAGGAAAGCAAGCGGAGGGGTGGCCCGGCTGGCTGGTGGCC 1664  
QY 1808 TCCGCTCTGTCTCGAGCCCTCTGTGGTCAAGCTGGATACAAAGATTCAGAGCCCTTCTCT 1867  
DB 1665 TCCGCTCTGTCTCGAGCCCTCTGTGGTCAAGCTGGATACAAAGATTCAGAGCCCTTCTCT 1724  
QY 1868 TGCTGTCTCAGCCGCTCAGGTTGGAGCCACAGACACCCAGCCAGCCCGGCTGGTCTG 1927  
DB 1725 TGCTGTCTCAGCCGCTCAGGTTGGAGCCACAGACACCCAGCCAGCCCGGCTGGTCTG 1784  
QY 1928 GGTCTCTCTCTGTGCTTTTCCCTCCAGAAATGCGGCTCAGAGCTTAGAAGCTCAACCCCC 1987  
DB 1785 GGTCTCTCTCTGTGCTTTTCCCTCCAGAAATGCGGCTCAGAGCTTAGAAGCTCAACCCCC 1844  
QY 1988 TATGAGGCGCAGCTCTCGGGTAGCTCTGAGCTCCGACCTCCGACCTTATGTCAAATTTTCAAC 2047  
DB 1845 TATGAGGCGCAGCTCTCGGGTAGCTCTGAGCTCCGACCTCCGACCTTATGTCAAATTTTCAAC 1904  
QY 2048 CATGGTTTTTCAATTTGACCCGCCCTTCTCGCTCATTAATGACACCCAGCTCTTTTGA 2107  
DB 1905 CATGGTTTTTCAATTTGACCCGCCCTTCTCGCTCATTAATGACACCCAGCTCTTTTGA 1964  
QY 2108 GGATCAGAGCCCATTTGCAACAGAGCGGCTGCGCAACCATCTTGTCTCTCGGATTTGCAA 2167  
DB 1965 GGATCAGAGCCCATTTGCAACAGAGCGGCTGCGCAACCATCTTGTCTCTCGGATTTGCAA 2024  
QY 2168 AATGACACCCAGGTAATCTAGAACATTTCTAAGGCCCTTTAACTCAGATGTCAAGCCACC 2227

Db 2025 AATGACACCCAGTAACTAGAAATCTCAAGCCCTTAACTCAGATGTCAAGCCACC 2084  
QY 2228 GGGCAAAACCCCGTCAATACCTCCACCAAGGAATGAGATATGGAACCTCACTGCTCCCC 2287  
Db 2085 GGGCAAAACCCCGTCAATACCTCCACCAAGGAATGAGATATGGAACCTCACTGCTCCCC 2144  
QY 2288 CAACCCAGGCTCAGGCTGGGACACGCAACGCTGTTCCGGGTTGGAACAGGAGGCTCA 2347  
Db 2145 CAACCCAGGCTCAGGCTGGGACATGCAACGCTGTTCCGGGTTGGAACAGGAGGCTCA 2204  
QY 2348 GAACTGGCTCTGAATAGCAGACCTAGCAGAGGAAGATACAGGGTATCGGGCGTTTG 2407  
Db 2205 GAACTGGCTCTGAATAGCAGACCTAGCAGAGGAAGATACAGGGTATCGGGCGTTTG 2264  
QY 2408 AGTGTTCAGAACTCAATTCGGGAAGATAAATCCAGTGCCTGGCCGACCACTGCAAT 2467  
Db 2265 AGTGTTCAGAACTCAATTCGGGAAGATAAATCCAGTGCCTGGCCGACCACTGCAAT 2324  
QY 2468 CAAGCTTGGACACAGGGGTTCTGTTCCGGAGGCAAAATTCCTTAGGAAAAAGAGACA 2527  
Db 2325 CAAGCTTGGACACAGGGGTTCTGTTCCGGAGGCAAAATTCCTTAGGAAAAAGAGACA 2384  
QY 2528 GACTTTCCTAAATGTGTCCAAATGCGGATCACTGCTCAGATGGAATCTAGAGCACTGAG 2587  
Db 2385 GACTTTCCTAAATGTGTCCAAATGCGGATCACTGCTCAGATGGAATCTAGAGCACTGAG 2444  
QY 2588 CTCCTGTCTCTGGAAGTATTTAAGAAAAAGGCTGGGCCAGGCAACGATGCTCACGCTGT 2647  
Db 2445 CTCCTGTCTCTGGAAGTATTTAAGAAAAAGGCTGGGCCAGGCAACGATGCTCACGCTGT 2504  
QY 2648 AATCCAGACTTTGGAGGCGCAGGCGAGGCGGATCACTGAGTGAGGAGTTTGAGAA 2707  
Db 2505 AATCCAGACTTTGGAGGCGCAGGCGAGGCGGATCACTGAGTGAGGAGTTTGAGAA 2564  
QY 2708 GCTGCGCAACATGGTGAACCTCATCTCTAATAAATAAATAAATAAATAAATAAATAA 2767  
Db 2565 GCTGCGCAACATGGTGAACCTCATCTCTAATAAATAAATAAATAAATAAATAAATAA 2624  
QY 2768 TGGCAGGTCCTGTATATCCAGCTACTTGGGAGGCTGAGGCATGAGATCACTTAAACCT 2827  
Db 2625 TGGCAGGTCCTGTATATCCAGCTACTTGGGAGGCTGAGGCATGAGATCACTTAAACCT 2684  
QY 2828 GAGAGCAGAGTTTACAGTGAGCCAGATCGTGCCCATGCAATTCAGGCTGGGCGACAGA 2887  
Db 2685 GAGAGCAGAGTTTACAGTGAGCCAGATCGTGCCCATGCAATTCAGGCTGGGCGACAGA 2744  
QY 2888 GCAAGACTCTGTCTC 2902  
Db 2745 GCAAGACTCTGTCTC 2759

## RESULT 2

US-11-072-512-986  
; Sequence 986, Application US/11072512  
; Publication No. US20060029945A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YUKI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAHOKO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOKYUKI  
; APPLICANT: NAGAHARI, KENJI

; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: Novel full length cDNA  
; FILE REFERENCE: 084335-0191  
; CURRENT APPLICATION NUMBER: US/11/072,512  
; CURRENT FILING DATE: 2005-03-07  
; PRIOR APPLICATION NUMBER: US 60/350,978  
; PRIOR FILING DATE: 2002-01-25  
; PRIOR APPLICATION NUMBER: JP 2001-379298  
; PRIOR FILING DATE: 2001-11-05  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 986  
; LENGTH: 2625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-072-512-986

Query Match 45.4%; Score 1325.8; DB 9; Length 2625;  
Best Local Similarity 99.8%; Pred. No. 5.1e-10;  
Matches 1338; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1562 CCCAGATCTCATCTGCTCATCTGAGTCCCAATCTTCCAGGGTGCAGCCCTCCG 1621  
Db 1286 CTCAGATCTCATCTGCTCATCTGAGTCCCAATCTTCCAGGGTGCAGCCCTCCG 1345  
QY 1622 TTCATCTCTGAAACCCAGCATCTTTTTCAGCTGCTTGAACATTTATTTTTTTT 1681  
Db 1346 TTCATCTCTGAAACCCAGCATCTTTTTCAGCTGCTTGAACATTTATTTTTTTT 1404  
QY 1682 AACGATGCAATTTTGTGCTTCCAGAAAGGGCCAGCTCTGAGCCCTCACCTTCCA 1741  
Db 1405 AACGATGCAATTTTGTGCTTCCAGAAAGGGCCAGCTCTGAGCCCTCACCTTCCA 1464  
QY 1742 CACTCAGAACTCTCAGCCGAGGAAGGCAAGAGCGCAGGGGTGGCGGTGCGTGC 1801  
Db 1465 CACTCAGAACTCTCAGCCGAGGAAGGCAAGAGCGCAGGGGTGGCGGTGCGTGC 1524  
QY 1802 GTGGCTCCGCTCTGCTCGAGCCCTGCTGAGCTGATACAGATTCAGATTCAGACCC 1861  
Db 1525 GTGGCTCCGCTCTGCTCGAGCCCTGCTGAGCTGATACAGATTCAGATTCAGACCC 1584  
QY 1862 TTCTCTGCTTGTACCCCGCTCCAGGTTGGAGCCACAGACACCCACCCCGGCTG 1921  
Db 1585 TTCTCTGCTTGTACCCCGCTCCAGGTTGGAGCCACAGACACCCACCCCGGCTG 1644  
QY 1922 GGTCTGCTCTCTCTGCTGCTTTCCTCCAGAAATGCGGCTCAGACTAGAACTCAA 1981  
Db 1645 GGTCTGCTCTCTCTGCTGCTTTCCTCCAGAAATGCGGCTCAGACTAGAACTCAA 1704  
QY 1982 CCCCCCTATGAGGGCCAGCTCTGGGGTAGCTCTGAGCTCCGACCTTATGTCCAAATTT 2041  
Db 1705 CCCCCCTATGAGGGCCAGCTCTGGGGTAGCTCTGAGCTCCGACCTTATGTCCAAATTT 1764  
QY 2042 CACACCATGGTTTTCATTTGACCCCGCCCTTCTGCTCATATATGACACCCAGCTCT 2101  
Db 1765 CACACCATGGTTTTCATTTGACCCCGCCCTTCTGCTCATATATGACACCCAGCTCT 1824  
QY 2102 TTGAGAGGATCAGAGCCCATTTGCAAGAAGAGCGGCTGCCAACCATCTTGTCTCCGA 2161  
Db 1825 TTGAGAGGATCAGAGCCCATTTGCAAGAAGAGCGGCTGCCAACCATCTTGTCTCCGA 1884  
QY 2162 TTGCAAAATGACCCCAAGTAACTAGAACATTTCAAGCCCTTTAACTCAGATGTCAA 2221  
Db 1885 TTGCAAAATGACCCCAAGTAACTAGAACATTTCAAGCCCTTTAACTCAGATGTCAA 1944  
QY 2222 GCCACCGGGCAAAACCCCGTCAATACCTCCCAAGGAATGAGATATGTGGACCTCATG 2281  
Db 1945 GCCACCGGGCAAAACCCCGTCAATACCTCCCAAGGAATGAGATATGTGGACCTCATG 2004  
QY 2282 CTCGCCCAACCCAGGCTCAGGCTGGGACACCGCCAAACGCTGTTCCGGGTGGAAACAGCAGA 2341  
Db 2005 CTCGCCCAACCCAGGCTCAGGCTGGGACACCGCCAAACGCTGTTCCGGGTGGAAACAGCAGA 2064



## RESULT 5

US-09-925-065A-594866  
; Sequence 594866, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 594866  
; LENGTH: 560  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-594866

Query Match 8.3%; Score 241.4; DB 6; Length 560;  
Best Local Similarity 85.1%; Pred. No. 7.6e+03;  
Matches 280; Conservative 1; Mismatches 47; Indels 1; Gaps 1;

QY 2590 CCTGTCTCTGGAAGTATTAAAGAAAGGCTGGGCCAGGACGATGGCTCAGCCCTGTAA 2649  
DB 1 CCACAGTGATGCAAGGTTCAAGAGTGTTCAGCCAGGCGATGGTGGCTCACACCTGTAA 60  
QY 2650 TCCAG-ACCTTGGAGGCGGAGGCGGATCCTCAGGTGAGGAGTTGAGACAG 2708  
DB 61 TCCAGACCTTGGAGGCGGAGGCTGGTGGATCCTCAGGTGAGGAGTTCAAGACAC 120  
QY 2709 CTGGCCAAATGCTGAAACCTCATCTCTACTAAATAAATAAATAAATAGCAGGCTGT 2768  
DB 121 CTGGCCAAATGAGAAACCCCTCTCTACTAAATAAATAAATAAATAGCAGGCTGT 180  
QY 2769 GGCAGTGGCTGTAAATCCAGCTACTTGGAGGCTGAGGATGAGAAATCACTTAAACCTG 2828  
DB 181 GGCAGTGGCTGTAAATCCAGCTACTTGGAGGCTGAGGATGAGAAATGCTTAAACCCA 240  
QY 2829 AGAGGCAGAGGTTACAGTGGCCCAAGATGCTGCTCAGCTGATTCAGCTGGGCGACAG 2888  
DB 241 GAGGTGGATGTTTCAAGTGGCCCAAGATGCTGCTCAGCTGATTCAGCTGGGCGACAG 300  
QY 2889 CAAGACTCTGTCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2917  
DB 301 CGAGATCTGTCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 329

## RESULT 6

US-10-330-773-159/c  
; Sequence 159, Application US/10330773  
; Publication No. US20060040262A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
; FILE REFERENCE: 529452001300  
; CURRENT APPLICATION NUMBER: US/10/330,773  
; CURRENT FILING DATE: 2002-12-27  
; NUMBER OF SEQ ID NOS: 981  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 159  
; LENGTH: 138627  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(138627)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-330-773-159

Query Match 8.3%; Score 240.8; DB 7; Length 138627;  
Best Local Similarity 89.1%; Pred. No. 83;  
Matches 271; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 2616 AGGCTGGCCAGGACGATGGCTCAGGCTGTAAATCCAG-ACCTTGGAGGCCGAGGCA 2674  
DB 108106 AGTGTGGCCAGGCGTGGCTCAGGCTGTAAATCCAGCACTTTGGAGGCCGAGGCG 108047  
QY 2675 GCGGATCACCTGAGGTGAGGAGTTTGAGAAACAGCCTGGCCCAACATGGTGAACCTCATC 2734  
DB 108046 GGTGATCATCTGCACTGAGGAGTTTCAGACGAGCTGCCCAAGATGGTGAACCTCATC 107987  
QY 2735 TCTACTAAATAATACAAAATTTAGCCAGGCTGGTGGAGGCTGCTGTATCCAGCTACT 2794  
DB 107986 TCTACTAAATAATTTCAAAAATTTAGCCAGGCTGGTGGAGGCTGCTGTATCCAGCTACT 107927  
QY 2795 TGGAGGCTGAGGCGATGAGATCACTTAAACCTGAGGCGAGGCTTACAGTGAGCCAG 2854  
DB 107926 CAGGAGGCTGAGGCGAGGAGATCGCTTGAACCTGGAGCGGAGGTTGAGTGAGCTAAG 107867  
QY 2855 ATCGTGCCACTGCATTCCAGCTGGCGGACAGAGCAAGACTCTGTCTCAAAAAAATAA 2914  
DB 107866 GTCATGCCACTGCCTCCAGCTGGCGGATAGAGCAAGCTGTCTCAAAAAAATAA 107807  
QY 2915 AAAA 2918  
DB 107806 AAAA 107803

## RESULT 7

US-09-925-065A-594865  
; Sequence 594865, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 594865  
; LENGTH: 560  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-594865

Query Match 8.2%; Score 239.8; DB 6; Length 560;  
Best Local Similarity 84.8%; Pred. No. 8e+03;  
Matches 279; Conservative 1; Mismatches 48; Indels 1; Gaps 1;

QY 2590 CCTGTCTCTGGAAGTATTAAAGAAAGGCTGGGCCAGGACGATGGCTCAGCCCTGTAA 2649

```
Db 1 CCCAGGTGCATGCAAGGTTCAAGAGTGGTTCCAGCCAGGCGATGGTGGCTCACACCTGTAA 60
QY 2650 TCCAG-ACITTTGGGAGCGGAGCGAGGCGGATCACCTGAGGTGAGGATTTGAGAACAG 2708
Db 61 TCCAGCACITTTGGGAGCGGAGGTTGGTGGATCCTGAGGTGAGGATTTCAAGACCAC 120
QY 2709 CTTGGCCCAACATGGTGAACCTCATCTCTACTAAAAATACAAAAATTTAGCCAGCGTGGT 2768
Db 121 CTTGGCCCAACATGGAGAACCCCATCTCTACTAAAAATACAAAAATTTAGCCAGCGTGGT 180
QY 2769 GCGAGTGCCTGTAAATCCAGCTTCTGGAGGCTGAGGATGAGATCACTTAAACCTG 2828
Db 181 GCGAGTGCCTGTAAATCCAGCTTCTGGAGGCTGAGGATGAGATTTCTTAAACCCA 240
QY 2829 AGAGGAGAGGTTACAGTACAGGATCGTGCCTGCACTCCAGCCTGGGCGACAGAG 2888
Db 241 GAGGTGGATGTTCTAGTACCCAGATCGTGCCTGCACTCCAGCCTGGGCGACAGAG 300
QY 2889 CAAGACTCTCTCTCAAAAAAATAAAAAA 2917
Db 301 CGAGATTCTCTCTCAAAAAAATAAAAAA 329
```

## RESULT 8

```
US-09-925-065A-55092
; Sequence 55092, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55092
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-55092
```

```
Query Match 8.2%; Score 239.6; DB 6; Length 628;
Best Local Similarity 87.3%; Pred. No. 7.3e+03;
Matches 274; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 2605 TATTTAAGAAAGGCTGGGCGAGCAGATGGCTCACGCTGTAATCCAG-ACITTTGGG 2663
Db 67 TAAAAAATAAAAAATTTGGCCGCGAGTGGCTCAACCTGTAATCCAGACITTTGG 126
QY 2664 AGCCAGGCGAGCGGATCACTGAGGTGAGGATTTGAGAAAGAGCTGGCCCAACATGTT 2723
Db 127 AGCCGAGGCGAGCAGATCACTTGAAGTCAGGAGTTGAGAGCCAGCTGGCCCAACATGTT 186
QY 2724 GAACCTCATCTCTACTAAAAATACAAAAATTTAGCCAGGCGTGGTGGCAGGCTGTAA 2783
Db 187 GAAACCTCATCTCTACTAAAAATACAAAAATTTAGCCGCGTGGTGGTGCACGCTGTAA 246
QY 2784 TCCAGACTACTTGGGAGGCTGAGGATGAGATCACTTAAACCTGAGAGCGAGGTTAC 2843
Db 247 TCCAGACTACTTGGGAGGCTGAGGAGGAGATCGCTTGAACCTAGAGGCGAGGTTGC 306
```

```
QY 2844 ACTGAGCCAGATCGTGGCCACTGCAATCCAGCTTGGGCGACAGAGCACTCTGTCTCA 2903
Db 307 ACTGAGCGAGATCGTGGCCACTGCAATCCAGCTTGGGCGACAGAGTGTGTCTCA 366
QY 2904 AAAAAAATAAAAAA 2917
Db 367 AAAAAAATAAAAAA 380
```

## RESULT 9

```
US-09-925-065A-482665
; Sequence 482665, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 482665
; LENGTH: 562
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-482665
```

```
Query Match 8.2%; Score 239.2; DB 6; Length 562;
Best Local Similarity 87.5%; Pred. No. 8.1e+03;
Matches 273; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

QY 2608 TTAAGAAAGGCTGGGCGAGCAGATGGCTCACGCTGTAATCCAG-ACITTTGGGAGG 2666
Db 159 TAAAGTATAATCTAGGCCAGGACCGTGGCTCATCTGCTGTAATCCAGAACTTTGGGGG 218
QY 2667 CCGAGGAGGCGGATCACTGAGGTGAGGATTTGAGAAAGAGCTGGCCCAACATGGTGA 2726
Db 219 CCGAGGCTGGTGGATCACCCTGAGGTGAAGAGATTGAGACCACCTCTGGCCCAACATGGTGA 278
QY 2727 ACCTCATCTCTACTAAAAATACAAAAATTTAGCCAGGCGTGGTGGCAGGTTGTAATCC 2786
Db 279 ACCCTGCTCTCTACTAAAAATACAAAAATTTAGCCAGGCGGCGTGGCAGGACCTGTAAATCC 338
QY 2787 CAGCTACTTGGGAGGCTGAGGATGAGATCACTTAAACCTGAGAGGCGAGGTTACAGT 2846
Db 339 CAGCTACTTGGGAAGCTGAGGCAAGAGATCACTTGAACCCGGGAGGCGAGGTTGCAGT 398
QY 2847 GAGCCAAAGATCGTGCCTGCACTGCACTTCCAGCTTGGGCGAGAGCAAGATCTGTCTCAAAA 2906
Db 399 GATCCAAAGATCATGTCCTGCACTTCCAGCTTGGGCGTGGTGGTGCAGAGTGAGACTCCGCTCTCAAAA 458
QY 2907 AAAAAAATAAAAAA 2918
Db 459 AAAAAAATAAAAAA 470
```

## RESULT 10

```
US-09-925-065A-563212/c
; Sequence 563212, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```



;; TITLE OF INVENTION: Identification and Mapping of Single  
;; Nucleotide Polymorphisms in the Human Genome  
;; FILE REFERENCE: 10827.135  
;; CURRENT APPLICATION NUMBER: US/09/925,065A  
;; PRIORITY FILING DATE: 2001-08-08  
;; PRIOR APPLICATION NUMBER: US 60/243,096  
;; PRIOR FILING DATE: 2000-10-24  
;; PRIOR APPLICATION NUMBER: US 60/252,147  
;; PRIOR FILING DATE: 2000-11-20  
;; PRIOR APPLICATION NUMBER: US 60/250,092  
;; PRIOR FILING DATE: 2000-11-30  
;; PRIOR APPLICATION NUMBER: US 60/261,766  
;; PRIOR FILING DATE: 2001-01-16  
;; PRIOR APPLICATION NUMBER: US 60/289,846  
;; PRIOR FILING DATE: 2001-05-09  
;; NUMBER OF SEQ ID NOS: 957086  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 563212  
;; LENGTH: 611  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-925-065A-563212

Query Match 8.2%; Score 239.2; DB 6; Length 611;  
Best Local Similarity 89.6%; Pred. No. 7.5e+03;  
Matches 267; Conservative 1; Mismatches 29; Indels 1; Gaps 1;  
QY 2622 GGCCAGGACGATGGCTCAGCGCTGTAATCCAG-ACCTTGGAGCGCCGAGCGCGGA 2680  
DB 468 GGCCAGGACCGTGGCTCATGTCTGTATYACAGCACTTTGGAGCGCGGTGGCGGA 409  
QY 2681 TCACCTGAGTGGAGGTTTGAGAACAGCCTGGCCAACTGTGAACTCTACT 2740  
DB 408 TCACCTGAGATCAGGAGTTTGAGACCAGCCTGGCCAACTGTGTAATCCAGCTACT 349  
QY 2741 AAAAAACAAAATAGCAGCGGTGGTGGCAGGTGCTGTATCCAGCTACTTGGGAG 2800  
DB 348 AAAAAACAAAATAGCAGCGGTGGTGGCAGGTGCTGTATCCAGCTACTTGGGAG 289  
QY 2801 GCTGAGGATGAGATCACTTAACTGAGGCGAGGTTACAGTGAGCCAGATCGT 2860  
DB 288 GCTGAGAGAGGAGATGCTTGAACCCAGGAGCAGAGTTGAGTCCAGATCGTG 229  
QY 2861 CCATGCACTCCAGCTGGCGGACAGCAAGCTCTGTCTCAAAAAA 2918  
DB 228 CCATGCACTCCAGCTGGCGGACAGCAAGCTCCGCTCAAAAAA 171

RESULT 11  
US-10-995-561-13342/c  
;; Sequence 13342, Application US/10995561  
;; Publication No. US2005027054A1  
;; GENERAL INFORMATION:  
;; APPLICANT: CARGILL, Michele et al.  
;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
;; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
;; FILE REFERENCE: CL001559  
;; CURRENT APPLICATION NUMBER: US/10/995,561  
;; CURRENT FILING DATE: 2004-11-24  
;; NUMBER OF SEQ ID NOS: 85702  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 13342  
;; LENGTH: 67126  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-995-561-13342

Query Match 8.2%; Score 239.2; DB 8; Length 67126;  
Best Local Similarity 86.4%; Pred. No. 1.6e+02;  
Matches 273; Conservative 2; Mismatches 40; Indels 1; Gaps 1;  
QY 2604 GTATTAAAGAGGCTGGCGGACGATGGCTCAGCGCTGTAATCCAG-ACCTTGG 2662

DB 46676 GAATTAAGTAAGTTTAGGGCTGGGCTCARTGGCTTATGCTGTAATCCAGCACTTTGG 46617  
QY 2663 GAGGCGGAGGCGGAGTCACTGAGGTGAGGAGTTTGAACAGCCTGGCCACATGG 2722  
DB 46616 GAGGCTGAGGCTGGGAGATCAGCTGAGGTCAAGAGTTGAGACTAGCCCTGGCCACATGG 46557  
QY 2723 TGAACCTCTCATCTCTACTAAATAACAAAAATTAGCCAGGCGTGGTGGCAGGTGCTGTGA 2782  
DB 46556 TGAAGGCGGCTCTCTACTTAAATAACAAAAATTAGCCAGGCGTGGTGGWAGGTGCTGTGA 46497  
QY 2783 ATCCAGCTACTTGGAGGCTGAGGCATGAGAACTCACTTAACTGAGAGGCGAGGTTA 2842  
DB 46496 ATCCAGCTACTTGGAGGCTGAGGCATGAGAACTCACTTAACTGAGAGGCGAGGCTG 46437  
QY 2843 CAGTGAGCAAGATCGTCCACTGCTATCCAGCTGGCGGACAGAGCAAGACTCTGTCTC 2902  
DB 46436 CAGCGAGCTGAGATCATGCCACTGCACTCCAGCCTGGATGACACAGCAAGACTCTGTCTC 46377  
QY 2903 AAAAAA 2918  
DB 46376 AAAAAA 46361

RESULT 12  
US-11-112-908-35/c  
;; Sequence 35, Application US/11112908  
;; Publication No. US20050260659A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Harris, Cole  
;; APPLICANT: Davis, Lisa M.  
;; TITLE OF INVENTION: Breast Cancer Biomarkers  
;; FILE REFERENCE: 04-164-US  
;; CURRENT APPLICATION NUMBER: US/11/112,908  
;; CURRENT FILING DATE: 2005-04-22  
;; PRIOR APPLICATION NUMBER: US 60/564,758  
;; PRIOR FILING DATE: 2004-04-23  
;; PRIOR APPLICATION NUMBER: US 60/575,978  
;; PRIOR FILING DATE: 2004-06-01  
;; PRIOR APPLICATION NUMBER: US 60/631,702  
;; PRIOR FILING DATE: 2004-11-30  
;; PRIOR APPLICATION NUMBER: US 60/633,826  
;; PRIOR FILING DATE: 2004-12-07  
;; NUMBER OF SEQ ID NOS: 511  
;; SOFTWARE: PatentIn version 3.3  
;; SEQ ID NO 35  
;; LENGTH: 127340  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-11-112-908-35

Query Match 8.2%; Score 239.2; DB 12; Length 127340;  
Best Local Similarity 87.5%; Pred. No. 92;  
Matches 273; Conservative 0; Mismatches 38; Indels 1; Gaps 1;  
QY 2608 TTAAGAAAGCTGGCGGAGCAGATGGCTCAGCGCTGTAATCCAG-ACCTTGGGAGG 2666  
DB 125149 TCAAAAAAAGGCGGCGGTGGCGCAGCGCTGTAATCCAGCACTTTGGGAGG 125090  
QY 2657 CCGAGGCGGCGGATCACTGAGGTGAGAGTTTGAAGAAAGCTGGCCAAATGGTCAA 2726  
DB 125089 CCGAGGCGGCGGATCACTGAGGTGAGAGTTTGAAGAAAGCTGGCCAAATGGCANA 125030  
QY 2727 ACCTCATCTCTACTTAAATAACAAAAATTAGCCAGGCGTGGTGGCAGGTGCTGTAAATCC 2786  
DB 125029 ACCCTGTCTCTACTTAAATAACAAAAATTAGCAGGCGTGGTGGCAGCACTGTAAATCC 124970  
QY 2787 CAGCTACTTGGAGGCTGAGCATGAGATCACTTAACTGAGAGGCGAGGTTTACAGT 2846  
DB 124969 CAGCTACTTGGAGGCTGAGCGGAGATCGCTTGAACCTTGAAGGCGAGTGGTTCAGT 124910  
QY 2847 GAGCAAGATCGTCCACTGCTTCCAGCCTGGCGGACAGAGCAAGACTCTGTCTCAAAA 2906

Db 124909 GAGCCAGATCATGCCACTCCAGCTCCAGCGCTGGGCAACAGAGTGAGACTCCATCTTAAAA 124850

Qy 2907 AAAAAAAAAAAAA 2918

Db 124849 AAAAAAAAAAAAA 124838

## RESULT 13

US-09-925-065A-621024/c

Sequence 621024, Application US/09925065A

Publication No. US20040181048A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

CURRENT FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20

PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

PRIOR FILING DATE: 2001-05-09

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 621024

LENGTH: 483

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-621024

## Query Match

Best Local Similarity 89.0%; Score 239; DB 6; Length 483;

Matches 268; Conservative 1; Mismatches 31; Indels 1; Gaps 1;

Qy 2619 CTGGGCGAGCAGATGGCTCAGCTGTAAATCCAG-ACCTTGGAGGCGGAGCGGC 2677

Db 483 CTGGGCGAGTGAGTGCGCTCATGCTGTAAATCCAGCACTTGGAGGCTGAGGCGGC 424

Qy 2678 GATACCTCAGGTGAGGAGTTTGAGAACAGCTCGCCCAACATGGTGAACCTCATCTCT 2737

Db 423 TGATCCTCAGGTGAGGAGTTGAGAACAGCTCGCCCAACATGGTGAACCTCATCTCT 364

Qy 2738 ACTAAATAACAAAAATTAGCCAGGCGTGTGCGAGGCTGTGTAAATCCAGCTACTTGG 2797

Db 363 ACTAAATAACAAAAATTAGCTGTGTGTGCGAGGCTGTGTAAATCCAGCTACTTGG 304

Qy 2798 GAGGCTGAGCATGAGATCACTTAACTTGAGGCGAGGTTACAGTGAGCCAGATC 2857

Db 303 GAGGCTGAGCATGAGATCACTTAACTTGAGGCGAGGTTACAGTGAGCCAGATC 244

Qy 2858 GTGCCACTGCATTCAGCGCTGGCGAGCAGAGCAAGACTCTGTCTCAAAAAA 2917

Db 243 GTGCCACTGCATTCAGCGCTGGCGAGCAGAGCAAGACTCTGTCTCAAAAAA 184

Qy 2918 A 2918

Db 183 A 183

## RESULT 14

US-09-925-065A-84578/c

Sequence 84578, Application US/09925065A

Publication No. US20040181048A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

CURRENT FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20

PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

PRIOR FILING DATE: 2001-05-09

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 84578

LENGTH: 1082

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-84578

## Query Match

Best Local Similarity 8.2%; Score 239; DB 6; Length 1082;

Matches 260; Conservative 1; Mismatches 36; Indels 0; Gaps 0;

Qy 2622 GGCAGGCGACGATGGCTCAGCGCTGTAAATCCAGACTTTGGGAGGCCAGGCGGAT 2681

Db 412 GGCAGGCGCGGTGGCTCAGTCTGTAAATCCAGCACTTGGGAGGCCAGGCGGAT 353

Qy 2682 CACCTGAGGTGAGGAGTTTGAGAACAGCTGGCCAACTGGTGAACCTCATCTCTACTA 2741

Db 352 CGCCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAGCATGGTGAACCTCATCTACTA 293

Qy 2742 AAAATACAAAAATTAGCCAGGCGGTGGCGAGGCTGTGTAAATCCAGCTACTTGGAGG 2801

Db 292 AAAATACAAAAATTAGCCAGGCGGTGGCGAGGCTGTGTAAATCCAGCTACTTGGAGG 233

Qy 2802 CTGAGGCGATGAGAACTCACTTAAACCTGAGAGCGAGGTTACAGTGAGCCAGATCGTGC 2861

Db 232 CTGAGGCGATGAGAACTCACTTAAACCTGAGAGCGAGGTTACAGTGAGCCAGATCGTGC 173

Qy 2862 CACTGCATTCAGCGCTGGCGAGCAGAGCAAGACTCTGTCTCAAAAAA 2918

Db 172 CATTGCACTCCAGCGCTGGCGAGGCTGTGTCTCAAAAAA 116

## RESULT 15

US-09-925-065A-723891/c

Sequence 723891, Application US/09925065A

Publication No. US20040181048A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

CURRENT FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20

PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

PRIOR FILING DATE: 2001-05-09

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 723891

LENGTH: 2649

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-723891

Query Match 8.2%; Score 238.6; DB 6; Length 2649;  
Best Local Similarity 88.5%; Pred. No. 2.3e+03;  
Matches 270; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

QY	2615	AAGCTGGGCGGCGAGGATGGCTCAGCGCTGTAATCCAG-ACCTTGGGAGGCCGAGGC	2673
DB	2260	AAAGATGGGCGGCGGCTGCTCAGCGCTGTAATCCAGCACTTTGGGAGGCCGAGGC	2201
QY	2674	AGCGGATCACCTGAGGTGAGGAGTTTGAGAACAGCCTGGCCAAACATGGTGAACCTCAT	2733
DB	2200	GGGTGGATCACTTGAGTCAAGGAGTTTGAGAACAGCCTGGCCAAACATGGTGAACCTCAT	2141
QY	2734	CTCTACTAAATAACAAAAATTAGCCAGGCGTGGTGGCAGGTGCCTGTAATCCAGCTAC	2793
DB	2140	CTCTACTAAAGATACAAAAATTAGCCAGGCTGGTGGTGGGCGCTGTAATCCAGCTAC	2081
QY	2794	TTGGGAGGCTGAGGCGATGAGAACTTAAACCTGAGAGCGACAGGTTACAGTGAGCCAA	2853
DB	2080	TCGGGAGACTGAGGCGAGGAGATCACTTGAACCTGGAGGCGAGGTTGCGTGAAGTGA	2021
QY	2854	GATCGTGCCACTGCATTCCAGCCTGGCGGACAGAGCAAGACTCTGTCTCAAAAAA	2913
DB	2020	GATTGTGACACTGCATCCAGCCTGGCGGACAAACAGACTCCATCTCAAAAAA	1961
QY	2914	AAAAA 2918	
DB	1960	AAAAA 1956	

Search completed: March 5, 2006, 07:06:18  
Job time : 815 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 03:31:06 ; Search time 11150 Seconds  
(without alignments)  
12244.368 Million cell updates/sec

Title: US-10-699-941-3

Perfect score: 2918

Sequence: 1 gcggcctctgccgcct.....tctcaaaaaaaaaa 2918

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hic:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gsa1:\*  
10: gb\_gsa2:\*  
11: gb\_gsa3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2782.6	95.4	3253	4	CR857898	Pongo pyg
2	940	32.2	1262	11	DQ053107	DQ053107 Homo sapi
3	857.8	29.4	2104	4	AK038483	AK038483 Mus muscu
4	857.8	29.4	2601	4	AK028298	AK028298 Mus muscu
5	857.8	29.4	2603	4	AK039411	AK039411 Mus muscu
6	857.8	29.4	3679	4	AK038990	AK038990 Mus muscu
7	831	28.5	2148	4	AK086723	AK086723 Mus muscu
8	783	26.8	996	11	DQ053108	DQ053108 Pan trogl
9	748	25.6	748	7	CN420201	CN420201 328775518
10	730	25.0	865	2	BG912726	BG912726 602807930
11	696.4	23.9	699	7	CV024962	CV024962 2512 Full1
12	693.4	23.8	714	8	DR002011	DR002011 TC120419
13	653	22.4	876	2	BG913524	BG913524 602811270
14	648.4	22.2	650	1	AL138250	AL138250 DKFp4547E
15	639	21.9	639	7	CN420203	CN420203 170006001
16	635.4	21.8	884	5	BUS21943	BUS21943 AGENCOURT
17	632.4	21.7	688	7	CR790459	CR790459 DKFp4549J
18	618	21.2	630	6	CB153985	CB153985 K-EST0211
19	595.6	20.4	842	2	BF345891	BF345891 602017783
20	585	20.0	609	2	BG910297	BG910297 602805869
21	578.6	19.8	800	6	CA315557	CA315557 UI-M-FW0
22	575.4	19.7	609	2	BG819886	BG819886 602781969

23	570.6	19.6	598	7	CR536713	CR536713 DKFp459E
24	565.6	19.4	606	1	AW961890	AW961890 BST373963
25	544.8	18.7	1022	2	BF530295	BF530295 602071474
26	544	18.6	1043	2	BG913501	BG913501 60281232
27	539	18.5	577	8	DR002579	DR002579 TC107021
28	538.2	18.4	689	3	BQ179996	BQ179996 UI-M-EW0
29	525.8	18.0	946	2	BF527108	BF527108 602039553
30	520.6	17.8	548	7	CR773543	CR773543 DKFp459E
31	502.6	17.2	673	7	CN219420	CN219420 WLA005H09
32	496.8	17.0	729	6	CA750371	CA750371 UI-M-FY0
33	483.4	16.6	509	7	CR628631	CR628631 DKFp459L
34	482.6	16.5	494	3	BQ312270	BQ312270 QV1-EN039
35	481.8	16.5	637	3	BQ044387	BQ044387 UI-M-EH0
36	459	15.7	608	2	BF527542	BF527542 602040445
37	441.4	15.1	1471	8	DN717620	DN717620 CNB150-CO
38	440.6	15.1	495	1	AI498192	AI498192 CNB119-F0
39	422.4	14.5	576	7	CN420200	CN420200 tm90912.x
40	410.6	14.1	521	6	CB720956	CB720956 AMGNNUC.N
41	409.8	14.0	1087	8	DN689249	DN689249 CGX63-E08
42	405.8	13.9	1402	8	DN723748	DN723748 CNB150-CO
43	390.2	13.4	752	6	CB245546	CB245546 UI-M-FY0
44	390.2	13.4	1169	8	DN735716	DN735716 CNB82-H02
45	386.4	13.2	461	1	AI480208	AI480208 tm72f02.x

## ALIGNMENTS

## RESULT 1

CR857898  
LOCUS  
DEFINITION  
Pongo pygmaeus mRNA; cDNA DKFp459H1854 (from clone DKFp459H1854).  
ACCESSION  
CR857898.1 GI:55726776  
VERSION  
CR857898.1  
KEYWORDS  
HTC.  
SOURCE  
Pongo pygmaeus (orangutan)  
ORGANISM  
Pongo pygmaeus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Pongo.  
REFERENCE  
1 (bases 1 to 3253)  
AUTHORS  
Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,  
Fobo, G., Han, M. and Wiemann, S.  
CONSRMT  
The German cDNA Consortium  
TITLE  
Direct Submission  
JOURNAL  
Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuherberg, GERMANY  
COMMENT  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by BMFZ (Biomedical Research Center at the  
Heinrich-Heine-University, Duesseidorf/Germany) within the CDNA  
sequencing consortium of the German Genome Project. This clone  
(DKFp459H1854) is available at the RZPD Deutsches  
Reisourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.  
Please contact RZPD for ordering:  
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFp459H1854  
Further information about the clone and the sequencing project is  
available at http://mips.gsf.de/projects/cdna/  
location/Qualifiers  
1. 3253  
/organism="Pongo pygmaeus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9600"  
/clone="DKFp459H1854"  
/tissue\_type="cortex"  
/clone\_lib="459 (synonym: pcorl). Vector pSport1\_Sfi; host  
DH10B; sites SfilA + SfilB"  
/dev\_stage="adult"  
/note="hypothetical protein (Homo sapiens)"  
1. 3253  
/gene="DKFp459H1854"  
408. .1523  
/gene="DKFp459H1854"

## FEATURES

source  
CR857898  
Pongo pygmaeus mRNA; cDNA DKFp459H1854 (from clone DKFp459H1854).  
ACCESSION  
CR857898.1 GI:55726776  
VERSION  
CR857898.1  
KEYWORDS  
HTC.  
SOURCE  
Pongo pygmaeus (orangutan)  
ORGANISM  
Pongo pygmaeus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Pongo.  
REFERENCE  
1 (bases 1 to 3253)  
AUTHORS  
Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,  
Fobo, G., Han, M. and Wiemann, S.  
CONSRMT  
The German cDNA Consortium  
TITLE  
Direct Submission  
JOURNAL  
Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuherberg, GERMANY  
COMMENT  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by BMFZ (Biomedical Research Center at the  
Heinrich-Heine-University, Duesseidorf/Germany) within the CDNA  
sequencing consortium of the German Genome Project. This clone  
(DKFp459H1854) is available at the RZPD Deutsches  
Reisourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.  
Please contact RZPD for ordering:  
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFp459H1854  
Further information about the clone and the sequencing project is  
available at http://mips.gsf.de/projects/cdna/  
location/Qualifiers  
1. 3253  
/organism="Pongo pygmaeus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9600"  
/clone="DKFp459H1854"  
/tissue\_type="cortex"  
/clone\_lib="459 (synonym: pcorl). Vector pSport1\_Sfi; host  
DH10B; sites SfilA + SfilB"  
/dev\_stage="adult"  
/note="hypothetical protein (Homo sapiens)"  
1. 3253  
/gene="DKFp459H1854"  
408. .1523  
/gene="DKFp459H1854"

```
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH90150.1"
/db_xref="GI:55726777"
/db_xref="GOA:Q5RDK7"
/db_xref="InterPro:IPR001251"
/db_xref="UniProt/TREMBL:Q5RDK7"
/translation="MGTTBATLRMENVDVKEWQDEDLPRPLPEETGVELLGSPVEDT
SSPNTLFNGAHRKRLTVPATKNMFGDSADLFGDGTEDGSAANGRLWRVTIG
EQRHLDLHWIRPKYKVTGCVYGEGLNLIIVPACTLPDSLSIDYHINENFLYV
ISLELLVAEDYIVYLNGATPRRMPGLSWLAKCTQMDRLRLKRLKSLIIVHPSWF
INTVLASPPFISVKPINKIQYVHSLDLEQLIPMEHVQIPDCVLOQYBEERLARRES
TRPQEFVLPRSEERPEVAPVENRSALVSEDOEISMS"
```

## ORIGIN

```
Query Match          95.4%; Score 2782.6; DB 4; Length 3253;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 2840; Conservative 0; Mismatches 74; Indels 3; Gaps 1;

QY 1 GCCAGCCTCTGCAGCCCTGAGCTGGGAAGACGAGCTACCTCGGAGGCGAGGCGCGCA 60
DB |
DB 36 GCCAGCCTCTGCCAGCCCTGAGCTGGGAAGACGAGCTACCTCGGAGGCGAGGCGCGCA 95
QY 61 GCGGCGCGCGATGAGAGGGGCGCAGCGCGCCCGCGCTGGGGAGCCACCGCTAAC 120
DB |
DB 96 GCGGCGCGCGATGAGAGGGGCGCAGCGCGCCCGAGCTGGGGAGCCACCGCTAAC 155
QY 121 CTGTGACCCACCCACCCCTGCACAAAGAGCTGGGGGCGCTGGCCAGTGCCTCGG 180
DB |
DB 156 CTGTGACCCACCCACCCCTGCACAAAGAGCTGGGGGCGCTGGCCAGTGCCTCGG 215
QY 181 TGAACCTTCTCGGATGAGAAATCCGCCCTCGGAGCATCTCTTCTCTAGGCTCTGAA 240
DB |
DB 216 TGAACCTTCTCGGATGAGAAATCCGCCCTCGGAGCATCTCTTCTCTAGGCTCTGAA 275
QY 241 GCGCGGGAGCGTGAGCGATGCCAGTGCACCCCGGCGAGGCTGCGCTTTGTTGCCA 300
DB |
DB 276 GCGCTGGGAGCGTGAGCGATGCCAGTGCACCCAGGCGAGGCTGCGCTTTGTTGCCA 335
QY 301 GTAAAGAGGAGGCTGCTCAGCTGCGAGGGGTATCCCTGCTTCAAGCCAGTGCCTC 360
DB |
DB 336 GTAAAGAGGAGGCTGCTCAGCTGCGAGGGGTATCCCTGCTTCAAGCCAGTGCCTC 395
QY 361 TTCCAGCTCCCATGGGACCAACGAGCAGCTCCGATGGAAACCTGACGCTGAAG 420
DB |
DB 396 TTCCAGCTCCCATGGGACCACTGAAGCAGCTCCGATGGAAACCTGACGCTGAAG 455
QY 421 GAGGAATGGCAGGACGAAATCTTCCAGGCCACTCCAGAGAGACGGGGGTGGAATG 480
DB |
DB 456 GAGGAATGGCAGGACGAAATCTTCCAGGCCACTCCAGAGAGACGGGGGTGGAATG 515
QY 481 CTGGCAGCCCGGTGGAAGACATCTCTCTCCCAACACGCTAAATTTCAACGGGCG 540
DB |
DB 516 CTGGCAGCCCGGTGGAAGACATCTCTCTCCCAACACACTAAATTTCAACGGGCG 575
QY 541 CATCGTAAGAGGAAGACGCTGTGGGCCCGCAGAGATCAACATTTCTCGGATCAGAGTGAG 600
DB |
DB 576 CATCGTAAGAGGAAGACGCTGTGGGCCCGCAGAGATCAACATTTCTCGGATCAGAGTGAG 635
QY 601 GGGTCCCTGTGTCGATGACTTTCTTTGGATACCCCTGATGACTCGGATATTAACGTGGAT 660
DB |
DB 636 GGGTCCCTGTGTCGATGACTTTCTTTGGATACCCCTGATGACTCGGATATTAACGTGGAT 695
QY 661 GACATCGAGACCCCGATGAGACTGACTCGTGGAGTCTCTGGGAATGGCAACGACTG 720
DB |
DB 696 GACATCGAGACCCCGATGAGACTGACTCGTGGAGTCTCTGGGAATGGCAACGACTG 755
QY 721 GAGTGGGAAGACGACACCCCGTGGCCACCGCCCAAGAACATGCCCGGGGACAGCGCGGAT 780
DB |
DB 756 GAGTGGGAAGATGACACCCCGTGGCCACTGGCCAGAACATGCCCGGGGACAGCGCGGAT 815
QY 781 CTATTGTTGGGACGGCAGCGAGGACGGCAGCGCCGCCAAACGGGGCGCTGTTGGCGGACA 840
```

```
DB |
QY 816 CTATTGTTGGGACGGCAGCGAGGACGGCAGTGCGCCCAACGGGCGCTGTTGGCGGACG 875
QY 841 GTGATCATCGGGGACGAGACGACCGTATAGACCTGACCATGATCGGCTTACATGAAA 900
DB |
DB 876 GTGATCATCGGGGACGAGACGACCGTATGACCTGCACTGATCCGCGCTTACATGAAA 935
QY 901 GTGGTCACCCACCGAGGAGTACTACGGCGAAGGCGCTCAACGCCCATCATCGTCTTCCAGCC 960
DB |
DB 936 GTGGTCACCCACCGAGGAGTACTATGGCGAAGGCGCTCAATGCCCATCATCGTCTTCCAGCC 995
QY 961 TGTCTCTTCCAGACAGAGCGCTCCCGCACTACCACTACATCATCATGAGAAACCTCTTCTG 1020
DB |
DB 996 TGTCTCTTCCAGACAGAGCGCTCCCGCACTACCACTACATCATGAGAAACCTCTTCTG 1055
QY 1021 TACGTCAATCAGAGCTTAGAGCTCCTGCTGCTGAGGACTACATCATCGTGTACTCTGAC 1080
DB |
DB 1056 TACGTCAATCAGAGCTTAGAGCTCCTGCTGCTGAGGACTACATCATCGTGTACTCTGAC 1115
QY 1081 GGTGCGACCGCCCGGCGGAGGATGCTGGAATCGGCTGGCTGGAAGAGTGTCTACCAAGATG 1140
DB |
DB 1116 GGTGCGACCGCCCGGCGGAGGATGCTGGAATCAGCTGGCTGAAGAGTGTCTACCAAGATG 1175
QY 1141 ATCGACCGGAGGTTGCGGAAAAAAGCTGAAAGTCTTTGATCATCGTCCACCCCTCGTGGTTC 1200
DB |
DB 1176 ATCGACCGGAGGTTGCGGAAAAAAGCTGAAAGTCTTTGATCATCGTCCACCCCTCGTGGTTC 1235
QY 1201 ATTCCGAGCTGTGCTGGCCCATCTCTCGCCCTTTTCATCAGGCTCAAGTTCATCAACAGATC 1260
DB |
DB 1236 ATTCCGAGCTGTGCTGGCCCATCTCTCGCCCTTTTCATCAGGCTCAAGTTCATCAACAGATC 1295
QY 1261 CAGTACGTGCGACAGCTTGGAAAGACCTGGAGCAACTCATCTTATGAAAACAGCTCCAGATC 1320
DB |
DB 1296 CAGTATGTGACAGCTTGGAGACCTGGAGCAACTCATCTCCCTATGGAACAGCTCCAGATC 1355
QY 1321 CCAGACTCGCTCTGCAATACGAGAGGAAAGACTGAAAGGCGAGAGGAGGAGCGGAGG 1380
DB |
DB 1356 CCAGACTCGCTCTGCAATACGAGAGGAAAGGCTGGAGGCGAGGAGGAGGAGCGGAGG 1415
QY 1381 CCCAGCGCGAGTGTGCTGCTGCCAGGCTTGAAGAGAGCCAGAGGTGGCACCAGTGGAA 1440
DB |
DB 1416 CCCAGCGCAGAGTGTGCTGCTGCCAGGCTTGAAGAGAGCCAGAGGTGGCACCAGTGGAA 1475
QY 1441 AACAGCTCTGCTCTGCTCTCAGAGATCAGGAAACAGCATGCTCTGAGGCGGACGTGAGC 1500
DB |
DB 1476 AACAGCTCTGCTCTGCTCTCAGAGATCAGGAAATAGCAATGCTCTGAGGCGGACGTGAGC 1535
QY 1501 ATAAACAAAGGACATGGAAGAAAGATTTCCAGATGCCAGAAAAACCTCTATCAGACGCCACTG 1560
DB |
DB 1536 ATAAACAAAGGACATGGAAGAAAGATTTCCAGATGCCAGAAAAACCTCTATCAGACGCCACTG 1595
QY 1561 GCCCAGATCTATCTGCTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
DB |
DB 1596 GCCCAGATCTATCTGCTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1655
QY 1621 GTTCACTCTGAAAACCCAGCATCTTTTTCAGCTGCTTGAATAACATTTGATTTTCTTTT 1680
DB |
DB 1656 GTTCACTCTGAAAACCCAGCATCTTTTTCAGCTGCTTGAATAACATTTGATTTTCTTTT 1712
QY 1681 TAAAGATGAGATTTTGTGCGTTCCAGAAAAAGGCGCCAGCTCTGAGCCCTCACCCCTTCC 1740
DB |
DB 1713 TAAAGATGAGATTTTGTGCGTTCCAGAAAAAGGCGCCAGCTCTGAGCCCTCACCCCTTCC 1772
QY 1741 ACACCTCAGCAACTCTCAGCCGAGGAGCAAGAGCGAGGGGTGGCCCGGCGGCGCTC 1800
DB |
DB 1773 ACACCTCAGCAACTCTCAGCCGAGGAGCAAGAGCGAGGGGTGGCCCGGCGGCGCTC 1832
QY 1801 GGTGGCTCTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
DB |
DB 1833 GGTGGCTCTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1892
QY 1861 CTTCTCTTGTGTCAACCCGCTCAGGTTGAGGCAACAGACCAACCGCCACCCCGGCT 1920
```









REFERENCE	1	Carninci, P. and Hayashizaki, Y.	
AUTHORS		High-efficiency full-length cDNA cloning	
TITLE		Meth. Enzymol. 303, 19-44 (1999)	
JOURNAL		10349636	
PUBMED			
REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
AUTHORS		Normalization and subtraction of cap-trapper-selected cDNAs to	
TITLE		prepare full-length cDNA libraries for rapid discovery of new genes	
JOURNAL		Genome Res. 10 (10), 1617-1630 (2000)	
PUBMED		11042159	
REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakauchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
AUTHORS		RIKEN integrated sequence analysis (RISA) system--384-format	
TITLE		sequencing pipeline with 384 multicapillary sequencer	
JOURNAL		Genome Res. 10 (11), 1757-1771 (2000)	
PUBMED		11076861	
REFERENCE	4	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.	
AUTHORS		Functional annotation of a full-length mouse cDNA collection	
TITLE		Nature 409, 685-690 (2001)	
JOURNAL			
PUBMED			
REFERENCE	5	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	
AUTHORS		Analysis of the mouse transcriptome based on functional annotation	
TITLE		of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
PUBMED			
REFERENCE	6	(bases 1 to 2601)	
AUTHORS		Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayaahida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tonari, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.	
TITLE		Direct Submission	
JOURNAL		Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of	
PUBMED		Physical and Chemical Research (RIKEN), Laboratory for Genome	
REFERENCE		Exploration Research Group, RIKEN Genomic Sciences Center (GSC),	
AUTHORS		RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,	
TITLE		Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,	
JOURNAL		URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,	
PUBMED		Fax: 81-45-503-9216)	
REFERENCE		cDNA library was prepared and sequenced in Mouse Genome	
AUTHORS		Encyclopedia Project of Genome Exploration Research Group in Riken	
TITLE		Genomic Sciences Center and Genome Science Laboratory in RIKEN.	
JOURNAL		Division of Experimental Animal Research in Riken contributed to	
PUBMED		prepare mouse tissues.	
REFERENCE		Please visit our web site for further details.	
AUTHORS		URL: http://genome.gsc.riken.jp/	
TITLE		URL: http://fantom.gsc.riken.jp/	
JOURNAL		Location/Qualifiers	
PUBMED		1. .2601	
REFERENCE		/organism="Mus musculus"	
AUTHORS		/mol_type="mRNA"	
TITLE		/strain="C57BL/6J"	
JOURNAL		/db_xref="FANTOM DB:3322401A10"	
PUBMED		/db_xref="taxon,10090"	
REFERENCE		/clone="3322401A10"	
AUTHORS		/tissue_type="head"	
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUB			

```
Db      747 GTGCGGACCTGTTTGGGGAGCGCTCTGCGGAAGACGCGATGCGGGCAACGCGTCTGT 806
QY      833 GGCGGACAGTGATCATCGGGGAGCAAGACGACCGGTATAGACCTGTCACATGATCGGCGCTT 892
Db      807 GGCGAATGTCATCATAGGGGAGCAAGACGATCGCATCGACCTGTCATATGATCGGCGCT 866
QY      893 ACATGAAGTGGTCAACCGAGGGTACTACCGGAGGCTTCAACGGCATCATGCT 952
Db      867 ACATGAAGTGGTCAACCGAGGATCTACCGGAGGCTTCAACGGCATCATGCT 926
QY      953 TCCGAGCTGCTTCTTCCAGACGAGCGCTCCCGACTACCACTACATCATGGAAC 1012
Db      927 TTGAGCGCTCTTCTGCGCAGATAGACGCTCCCGACTATCATCATGAGGAATC 986
QY      1013 TCTTCTGTAGCTCATCAGCAGCTTAGAGCTCTCTGCTGGCTGAGGACTACATCATGCTGT 1072
Db      987 TCTTCTGTAGCTCATCAGCAGCTTAGAGCTCTCTGCTGGCTGAGGACTACATCATGCTGT 1046
QY      1073 ACTGAAGCTGTCACGCGCGCGGAGATGCTGGAATCGCTGGCTGAAGAGTGT 1132
Db      1047 ATCTGAAGCGCGCGCGCGGAGGATGCTGCGCATTTGGCTGAAGAGTGT 1106
QY      1133 ACCAGATGATCGACCGGAGTTCGCGAAGAACTGAGTCTCTGATCATGCTCCACCT 1192
Db      1107 ACCAGATGATTCAGGAGACTGAGGAAGAACTTAAAGTCCCTGATCATGCTCCACCT 1166
QY      1193 CGTGGTTCATTCGAGCTGTGCTGCGCATCTCTGCGCTTTCATCGCTCAAGTTCATCA 1252
Db      1167 CTGTTTCATTCGATCTGTGCTGCGCATCTCTGCGCTTTCATCGCTCAAGTTCATCA 1226
QY      1253 ACAAGATCCAGTACGTCAGCTTGGAGAGCTCGGAGAACTCATCTCTGGAACAG 1312
Db      1227 GTAAATTCAGTACGTCAGCTTGGAGAGCTCGGAGAGCTGATTCCTCATGGAACAG 1286
QY      1313 TCCAGATCCAGTACGTCAGCTTGGAGAGCTCGGAGAACTCATGAGGAGGAGGAGA 1372
Db      1287 TGCAGCTGCGAGACTGTGCTGCTCAATATGAAGAGCAGAGACTCCGAGCAGAGGAG 1346
QY      1373 GGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1429
Db      1347 GCACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1406
QY      1430 CACAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1489
Db      1407 TGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1466
QY      1490 GCGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1549
Db      1467 CTACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1523
QY      1550 GAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1586
Db      1524 GAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560

RESULT 5
AK039411
LOCUS   AK039411.1 GI:26333248
DEFINITION Mus musculus adult male spinal cord cDNA, RIKEN full-length
           enriched library, clone:A330040P09 product:weakly similar to
           BCL2/ADENOVIRUS E1B 19-KDA PROTEIN-INTERACTING PROTEIN 2 [Mus
           musculus], full insert sequence.
ACCESSION AK039411
VERSION   AK039411.1
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
           Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
           Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS   Carninci, P. and Hayashizaki, Y.
```

TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORSTITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORSTITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORSTITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORSTITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORSTITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS

## COMMENT

FEATURES  
source

1. 2603  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM,DB:A330040P09"  
/db\_xref="taxon:10090"  
/clone="A330040P09"  
/sex="male"  
/tissue type="spinal cord"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"

High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
10349636

2  
Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
11042159

3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, K.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
11076861

4  
The RIKEN Genome Exploration Research Group Phase II Team and the  
PANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

5  
The PANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

6 (bases 1 to 2603)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,  
Hori, P., Imotani, K., Iehli, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koye, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.

Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,  
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

location/Qualifiers

1. 2603

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="PANTOM,DB:A330040P09"

/db\_xref="taxon:10090"

/clone="A330040P09"

/sex="male"

/tissue type="spinal cord"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"



CDS  
/dev\_stage="adult"  
348..1466  
/notes="unnamed protein product; putative  
weakly similar to BCL2/ADENOVIRUS E1B 19-KDA  
PROTEIN-INTERACTING PROTEIN 2 [Mus musculus]  
(SWISSPROT|O54940, evidence: PASTY, 52.5%ID, 94.7%length,  
match=972)"  
/codon\_start=1  
/protein\_id="BAC30342.1"  
/db\_xref="GI:26333249"  
translation="MCTEATLMEVNDVDEWDLRPLPDPEDTGVRLGAVEDS  
SPSPSTLNSGARKKTLVAPENISLQSEGLSDFLDTPDDLDINDDIETPD  
ETDLSFLNGNELEWEDDTPVATKMPGDSADLFGDSADGSAANRLWRITLIG  
EQBRHDLIRMYRVTHGGYVGGELNAIIVFAACFLPDS SPDYHYIMENLFLYV  
ISLELLVAEDYVIVLNGATPRRRMPGIGLWKYHMDRLRKLKSLIIVHPSWF  
IRTVLAISRPFSIVTSKIQVHSLERLERLIPMEHVOLPDCVLQVQRRLAKRES  
TRPQPEFLPRSEKPEIVBEDRAAEATEDQETSMS"

ORIGIN

Query Match 29.4%; Score 857.8; DB 4; Length 2603;  
Best Local Similarity 75.1%; Pred. No. 3.7e-121;  
Matches 1155; Conservative 0; Mismatches 337; Indels 45; Gaps 5;  
QY 64 GGGGGGGATGACAGGGGGCGACCGCGCGCTGGGGAGCCACCGCTAACCT 123  
DB 56 GGGCTCGGATGGCAGGGGAAACAGTGACTAAAGGGGGACAGGGGGTCTGCTACTAACCCA 115  
QY 124 GCAACCCACCCACCTCGACAAAGAGCTGGCGGGCGCTGGCCAGCTCGCCCTGGGTGA 183  
DB 116 CGCGGGCGCTTCTACAGTTGGGCGCGACAGGTGTCTGTGGCCACGTGCCCTGGGTGA 175  
QY 184 CTTCTCTCGGATCGAAGTCCGCCCTCGAGCATCTCTCTCTCTAGGCTCTGAAGGC 243  
DB 176 CTTCTCTCAGATGTGA-----CTTGGCGCTGAGCAT 207  
QY 244 CGGGGAGCGTGAGCGATGCCAGCTGCACCGGGCAGGGCTCGCTCTTTGTTGCCAGTA 303  
DB 208 CTTTCCACAGGGCCCTCGCGCTGGGTACCATCGAGAAATGCCCGCCCTTTGTGTCCAAGTG 267  
QY 304 AG-GAGGAGAGGCTGTCTAGCTGCAGAG-----GGGTATCCCTGCTTCAAGCC 352  
DB 268 ACAGCGCAGAGGCGAGCTTCGCTACCGAGATCATCTTCTGGGTACCCCGAGTTTCAGAC 327  
QY 353 AGTGCCTTTTCCAGCTCCCATGGGACCAACGAGCCACGCTCCGGATGGAAAACGTTGG 412  
DB 328 CACTTCTCTTTCAGCTCTCATGGAAACACAGAACTACACTAAGGATGGAAATGTGG 387  
QY 413 ACGTGAAGGGAATGGCAGGACGAAGATCTTCCAGGCCAATCCAGAGAGACGGGG 472  
DB 388 ACGTGAGGGATGAATGGCAGGATGAGGATCTGCCACAGCGCTCCCAAGAGACACCGGG 447  
QY 473 TGGAACTGCTTGGCAGCCGGTGGAAAGACATCTCTCTCTCCCAACAGCTAAATTTCA 532  
DB 448 TGGAGCGGCTGGGTGGCGAGTGGAAAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 507  
QY 533 ACGGAGCGCATCTGAAGAGAGACGCTGGTGGCCCGCCAGAGATCAACATTTCTCTGGATC 592  
DB 508 GCGGAGCACAATCGAAAGAGAAAGACGCTGGTGGCTCCAGAGATCAACATCTCCCTGGACC 567  
QY 593 AGAGTGAAGGGTCCCTGCTGCTCGATGACTTCTTGATACCTCTGATGACCTGGATATTA 652  
DB 568 AAAGCAGGGGCTCTCTGCTGTCCGACGACTTCTCGACACACCTGTGATGACCTGGACATCA 627  
QY 653 ACCTGGATGACATCGAGACCCCGCGATGAGACCGACTCGCTGGAGTTTCTGGGGAATGGCA 712  
DB 628 ATGTGAGCAGCATTTGAGACCGCAGATGAAGTGAATCTCTCTGGAGTTCTTGGGAATGGCA 687  
QY 713 ACAGACTGGAGTGGGAGACGACACCCCGTGGCCACCGCCAGAGACATGCCCGGGGACA 772  
DB 688 ATGAACTTGAGTGGGAGATGACACCCAGTGGCCACCGCCAAAGACATGCTGTGTGACA 747  
QY 773 GCGCGGATCTATTGGGAGCGGACGACCGAGAGACGGCAGCGCGCCCAACGGCGCTGT 832

DB 748 GTGGGACCTGTGTTGGGACGGCTCTGCGAAGACGGCAGTGGGCCAACGGTCTGTCT 807  
QY 833 GCGGACAGTGTATCATCGGGAGCAAGAGACCGGTATAGACCTGCACATGATCCGSCCTT 892  
DB 808 GCGGACAGTGTATCATAGGGAGCAAGAGCATGCGATCGACCTGATGATGATCCGSCCT 867  
QY 893 ACATGAAGTGTGTACCCACGAGGGGTACTACGGGAAGGGCTCAACGCGCATCATCGTCT 952  
DB 868 ACATGAAGTGTGTACCCATGAGGATACTACGGGGAAGGTCTCAACGCGCATCATCGTCT 927  
QY 953 TCGAGCGCTGCTCTCTTCAGACACAGCGCTCCCGACCTACCATCATCATGAGGAACC 1012  
DB 928 TTGAGCGCTGCTCTCTTCGAGATAGCAGCTCCCGACATCATCATCATCATGAGGAATC 987  
QY 1013 TCTCTCTGTACGTATCATGACAGCTTAGAGCTCTCTGCTGGCTGAGGACTACATGATCGTCT 1072  
DB 988 TCTCTCTGTACGTATCATGACAGCTTAGAACCTCTCTGCTGGGCTGAGGACTACATGATCGTCT 1047  
QY 1073 ACCTGAACGGTGTCCACGCGCGCGGAGGATGCTCTGGAATCGGCTGGCTGAAGAAGTGTCT 1132  
DB 1048 ATCTGAACGGGCGCACGCGCGCGGAGGAGTGGCTGGCATTTGGTGGCTGAAGAAGTGT 1107  
QY 1133 ACCAGATGATCGACCGGAGGTTTGGGAAAAACCTGAAGTCTTGTATCATCATCGTCCACCT 1192  
DB 1108 ACCACATGATTTGACAGGAGACTGAGGAAGAATCTAAAGTCCCTGATCATCTGCTCCACCT 1167  
QY 1193 CQTGGTTCATTGGGAGTGTCTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTTCATCA 1252  
DB 1168 CQTGGTTCATTGGGAGTGTCTGGCCATCTCCCGCCATTCATCAGTGTCAAGTTTCATCA 1227  
QY 1253 ACAAGATCCAGTACGTGCACAGCTTTGGAAGACTGGAGCACTCATCCCTATGGAACAG 1312  
DB 1228 GTAAATTCAGTACGTGCACAGCTTGAAGAGCTGGAGGAGCTGATTTCCATGGAACAG 1287  
QY 1313 TCCAGATCCAGACTGCTCTGCAATAGCAAGAGAAAGACTGAAGCCAGGAGGAGAGA 1372  
DB 1288 TGACAGTCCAGACTGTCTCTGCAATATGAAGAGCAGAGACTCCGAGGCAAGAGGAGAGA 1347  
QY 1373 GCGCGAGG---CCCCAGCGCGAGTTTGTCTGCCCGAGTCTCAAGAGAGCCAGAGGTGG 1429  
DB 1348 GCAACGCGCCACCGCGCGGAGTTCTCTCTTCCAGGTTCAGAGAAAGCCAGAGACTG 1407  
QY 1430 CACCACTGGAAAAAGGCTGTCTGTCTCTCAGAAAGATCAGAAAAACAAGCATGTCTGAG 1489  
DB 1408 TGAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1467  
QY 1490 GCGAGTGAAGCATAAAGAGGACATGGAAGAGATTCAGATGCCAGAAAAACCTCTGTCA 1549  
DB 1468 CTACCCAGAACCTAGACATGGACA---GAGATTTATTTCCCAACATCATGTAAACCTCATCC 1524  
QY 1550 GAGCGCCACTGGCCCGCAGATCTCATCTCTGCTGCTCATCC 1586  
DB 1525 GAAGCCCTGGGAGCCCGGCTTTCATCGAGGCTTCCACC 1561

RESULT 6  
AK038990

LOCUS  
DEFINITION

AK038990 3679 bp mRNA linear HTC 03-APR-2004  
Mus musculus adult male hypothalamus cDNA, RIKEN full-length  
enriched library, clone:A230081007 product:weakly similar to  
BCL2/ADENOVIRUS E1B 19-KDA PROTEIN-INTERACTING PROTEIN 2 [Mus  
musculus], full insert sequence.

ACCESSION

AK038990.1 GI:26332954

VERSION

HTC; CAP trapper.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning



## JOURNAL

PUBMED  
10349636

## REFERENCE

2

## AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)

11042159

## JOURNAL

3

## REFERENCE

4

## AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kikunai, T., Tashiro, H., Itoh, M., Sumi, N., Iehi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

11076861

## JOURNAL

4

## REFERENCE

5

## AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium

## TITLE

Functional annotation of a full-length mouse cDNA collection

## JOURNAL

5

## REFERENCE

6

## AUTHORS

Nature 409, 685-690 (2001)  
The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 3679)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Haeizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Horii, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T., Katoh, H., Kawai, J., Kojima, S., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Takahashi, P., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasumishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission

## JOURNAL

6

## REFERENCE

7

## AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.jp/  
URL: http://fantom.gsc.riken.jp/

## FEATURES

source

1. .3679  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM DB:A230081007"  
/db\_xref="taxon:10090"  
/clone="A230081007"  
/sex="male"  
/tissue\_type="hypothalamus"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"

## CDS

341. .1459

/note="unnamed protein product; putative weakly similar to BCL2/ADENOVIRUS E1B 19-KDA PROTEIN-INTERACTING PROTEIN 2 [Mus musculus] (SWISSPROT|054940, evidence: PASTY, 52.5%ID, 94.7%length, match=972)"  
/codon\_start=1  
/protein\_id="BAC30195.1"  
/db\_xref="GI:26332955"

/translation="MGTTAETLRMENVDVDEWDEDLPRLPEDTGVRLGGAVEDS

SSPSTNLNNGRRKRLVAPETINISLDQSEGLSDPFLDTPDDLDIVDDIETPD  
ETDSDFGNGLEWEDTPVATKNNPGLSADLFGDSABDGSAAAGRLWRVTVIG  
BOEHRDLHMIRPYMKVYHGGYEGEGLNAIVFAACFLPSSSDYHYIMENFLYV  
ISSLELVAEDYMIYVLNAGPRRPMGIGLKKCYHMDRLRLKSLIIVHPSWF  
IRTVLAISRPFISVKFISKIOVHSLERLELIPMHRVQLPCVLQVEQRRLAKRES  
TRPPQPELLRPERSEKPTVEEDRAAETEDQETMS"

3659. .3664

/note="putative"

3679

/note="putative"

## ORIGIN

Query Match	Best Local Similarity	Score	DB 4;	Length
64	GGGCGCGATGAGAGGGGGCGCGAGCCCGCGCTGGGAGCCACCGCTAACCT	123		
49	GGGCTGCCATGGCAGGGGAAACAGTGACTAAAGGGGACAGGGGTGCTACTAACCA	108		
124	GCACCCACCCACCCCTGCACAAAGAGCTGGCGGGCGCTGGCCACAGCTCGCCCTGGTGA	183		
109	CGGCGCGCTTCTACAGTGGCGCGGACAGGTGTGTGGCCAGCTCGCCCTGGTGA	168		
184	CTTCTCTCGATGACAGATCCGCCCTCGGAGCATCTCTCTCTCTAGGCTCTCAAGGC	243		
169	CTTCTCTCAGATGGGA-----CTTGGCCTCGAGCAT	200		
244	CGGGGAGCGTGAGCGATGCCCGCTGCACCGGCGAGGGCTCGCTTGTTCAGTA	303		
201	CTTCCACAGCGCCCTCGCTGGGTACCATCGAAGATCCCGCGCTTGTTCAGTA	260		
304	AG-GAGGAGAGCTGTCTCAGCTGCAGAG-----GGGTATCTCTCTCTCAAGCC	352		
261	ACAGCGCAGAGCGAGCTTCGCTACCGAGATCATCTTCTGGTCACTCCAGTTTCAGAC	320		
353	AGTGCTCTTCCAGCTCCCATGGGACACCGAAGCCACCGCTCCGATGGAAACGTGG	412		
321	CACCTTCTTTCCAGCTCTCATGGGAACACAGAGCTACACTAAGGATGGAAATGTGG	380		
413	ACGTGAAGGAGGATGGCAGACGAGATCTTCCAGGCCACTCCAGAGAGACGGGG	472		
381	ACGTGAGGAGGATGGCAGGATGAGGATCTGCCAGACCGCTCCAGAGACACCGGG	440		
473	TGGAATCTCTGGAGCGCGGTGGAGACACATCTCTCTCCCAACACGCTTAATTTCA	532		
441	TGGAGCGCTGGTGGCGCAGTGGAGAGCTCTCTCTACCTCTCTCCCTGAACTTGA	500		
533	ACGAGCGCATCGTAAGAGGAGAGCGTGGTGGCCCGCAGAGATCAACATTTCTCTGGATC	592		
501	CGGAGGACATCGAAAGAGAGAGCGTGGTGGCTCCAGAGATCAACATCTCCCTGGAC	560		
593	AGAGTGGGGTCTCTGTCTCGATGATCTTCTGGATGATCCCTGATGACCTGGATATTA	652		
561	AAAGCGAGGCTCTCTGTCTCGAGCTTCTCTCGACACACCTGATGACCTGGACATCA	620		
653	ACGTGGATGATCGAGACCGCCCGATGAGACCGAGTCTCTGGAGTCTCTGGGATGCA	712		
621	ATGTGACGACATTTGAGACCGCAGATGAATGACTCTCTGGAGTCTTGGGAATGCA	680		
713	ACGAACTGGAGTGGGAAGACGACACCCCGTGGCCACCGCCCAAGAACATGCCCGGGACA	772		
691	ATGAACTGGAGTGGGAAGATGACACCCCGAGTGGCCACCGCCCAAGAACATGCTGTGACA	740		



/clone\_lib=RIKEN full-length enriched mouse cDNA library  
/dev\_stage="15 days embryo"  
1. 2148  
/notes="weakly similar to BCL2/ADENOVIRUS E1B 19-KDA  
PROTEIN-INTERACTING PROTEIN 2 [Mus musculus]  
(SWISSPROT|054940, evidence: PASTY, 52.5%ID, 94.7%length,  
match=972)"

## misc\_feature

## ORIGIN

Query Match 28.5%; Score 831; DB 4; Length 2148;  
Best Local Similarity 74.9%; Pred. No. 4.7e-117; Indels 47; Gaps 7;  
Matches 1152; Conservative 0; Mismatches 340;  
64 GGGCGGCGATGAGAGGGGGCGAGCCGAGCCCGCTGGGAGGCCACCGCTAACCCCT 123  
DB GGGCTCGGATGGCAGGGGAAACAGTGACTAAAGGGGACAGGGGTGCTGCTACTAACCCA 148  
QY 124 GCACCCACCCACCCCTGACAAAGAGCTGGGGGGCTGGCCAGCTGGCCCTGGGTGA 183  
DB 149 CGGCGCGGCTTCTACAGTTGGGCGGACAGGTGTCTGTGGCCACGCTGGCCCTGGGTGA 208  
QY 184 CTTTCTCGGATGACAGATCCGCGCCCTGGGAGCATCTCTTCTCTAGGCTCTGAAGGC 243  
DB 209 CTTTCTCAGATGTGA-----CTTGGCCCTGAGCAT 240  
QY 244 CGGGGAGCGGTGAGCGATGCCAGCTGCAACCGGCGAGGGCTGCGCTTTGTTTGGCAGTA 303  
DB 241 CTTTCCACAGGCGCCCTCGCTGGGTACCATCGAGAAATGCCCGCTTTGTGTCCAAGTG 300  
QY 304 AG-GAGGAGGCTGTCTCAGCTGACAG-----GGGTATCCCTGCTTCAAGCC 352  
DB 301 ACAGCGCAGGCGGAGCTTCGGCTACCGAGATCATCTTCTGGGTCAACCGGATTTCCAGAC 360  
QY 353 AGTGCCTTCCAGCTCCCATGGGA-CCACGAGACCAAGCTCCCGGATGGAACGTTG 411  
DB 361 CACTTCTTTCCAGCTCTCATGGAAACCAAGAGCTACATAAGGATGGAATATG 420  
QY 412 GACGTGAAGGAGGAATGGCAGGACGAGATCTTCCAGGCCCATCTCCAGAGAGACGGGG 471  
DB 421 GACGTGAGGATGAAATGGCAGGATGAGGATCTGCCAGACCGCTCCAGAGACACCGGG 480  
QY 472 GTGGAACTCTGGGAGCGCGGTGGGAAGACATCTCTCTCCGCAACAGCTTAATTC 531  
DB 481 GTGGAGCGCTGGGTGGCGCAGTGGGAAGACTCTCTCTCACTCCCTCCACCTGAACTTG 540  
QY 532 AACGAGCGCATGTAAGAGAGAGCGCTGTGGCCCGCAGAGATCAACATTTCTTGAT 591  
DB 541 ACGGAGCACATCGAAAGAGAGACGCTGGTGGCTCCAGAGATCAACATCTCCCTGGAC 600  
QY 592 CAGAGTGAAGGGTCCCTGCTGCCGATGACTTCTTGATACCCCTGATGACCTGGATAT 651  
DB 601 CAAAGCGAGGGCTCTCTGCTGCCGAGACTTCTCGACACACCTGATGACCTGACATC 660  
QY 652 AACGTGGATGACATCGAGACCCCGGATGAGACCGACTCTCGTGGATTCCTGGGGAATGC 711  
DB 661 AATGTGGACGACATTCAGAGCGCCAGATGAACTGACTCTCTGGATTTCTGGGAAATGC 720  
QY 712 AACGAACTGAGTGGGAAGACGACACCCCGCTGGCCACCGCAAGAGATGCCGGGAC 771  
DB 721 AATGAATCTGAGTGGGAAGATGACACCCAGTGGCCACCGCAAAACATGCTGCTGTGAC 780  
QY 772 AGCGCGGATCTATTGGGGA-CGGCACGACGAGGAGCGGCGCGCCCAACGGGCGCT 830  
DB 781 AGTGGGACCTGTTTGGGACCGGCTCTCGGAGAGCGGAGTGGCGCCACAGCTGCTCT 840  
QY 831 GTGGCGGACAGTGATCATCGGGAGCAAGACACCGTATAGACTCTGCATGATCGGCC 890  
DB 841 GTGGCGAATGTCATCATAGGGAGCAAGAGCATCGCATCGACCTGCATGATCGGCC 900  
QY 891 TTACATGAAGTGTCAACCCAGGAGGTACTACGGGAGAGCGCTCAAGCCCATCATGCT 950  
DB 901 CTACATGAAGTGTCAACCCATGGAGTACTACGGGGAAGTCTCAAGCCCATCATGCT 960

QY 951 CTTCCAGCGCTGCTTCTTCTTCCAGACAGCAGCGCTCCCGACTACCACTACATCATGAGAA 1010  
DB GTTTGCGAGCTGCTTCTTCTTCCAGATAGCAGCTCCCGAGACTATCACTACATCATGAGAA 1020  
QY 1011 CTTCTTCTTGTAGCTCATCAGCAGCTTGTAGCTCTTGTGGTGTGAGGACTACATGATCT 1070  
DB TCTTCTTCTTATACGTCATCATCAGCTAGAACTGCTGTGGTGTGAGGACTACATGATCT 1080  
QY 1071 GTACCTGAAGCGTGCACGCGCGCGGAGGATGCTGGAATCGGCTGCTGAAGAGTG 1130  
DB GTATCTGAAGCGCGCACGCGCGGAGGAGGATGCTGCAATGCTGCTGCTGCTGCTGCTGCT 1140  
QY 1131 CTACAGATGATTCGACCGGAGGTTTCGGAATAAAGCTGAAAGTCTTGTATCATCTGTCACCC 1190  
DB TTACACATGATTGACAGGAGACTGAGGAGATCTTAAGTCCCTGATCATCTGTCACCC 1200  
QY 1191 CTCGTGGTTTCTTGGAGCTGTGCTGGCCATCTCTCGCCCTTTCATCAGCGTCAAGTTCT 1250  
DB CTCCTGGTTTCTTGGCACTGTGCTGGCCATCTTCCCGGCAATTCATCAGTGTCAAGTTCT 1260  
QY 1251 CAACAAGATCCAGTACGTGCAAGCTTGGAGACCTGAGGCAACTCATCCCTATGGAACA 1310  
DB CAGTAAATTCAGTACGTGCAAGCTTGGAGAGCTGAGGAGCTGATTTCCATGGAACA 1320  
QY 1311 CTTCCAGATCCAGACTGCTGCTCAATACAGAGAGAAAGACTGAAAGCCAGGAGGA 1370  
DB CTTGCGAGCTGCGAGACTGTGCTCTCAATATGAGAGCAGAGACTCCGAGCCAGGAG 1380  
QY 1371 GAGCGCGAGG---CCCCAGCGGAGTTTGTGCTGCCAGGTCTGAAGAGAACCCAGAGGT 1427  
DB GAGCACACGCGGCACGCGCAGCGAGTCTCTCTTCCAGGTCTGAGAGAAAGCCAGAGAC 1440  
QY 1428 GGCACCACTGGAAGACAGCTGCTGCTGCTGCTCAGAGATCAGGAAACAGCATGCTCTG 1487  
DB TGTGGAAGAGAGGACAGGCGAGGAGGCAACAGGAGCCAGGAAACTAGCATGCTCTG 1500  
QY 1488 AGCGGACGTGAGCATTAACAAAGGACATGGAAGAGATCTCCAGATGCCAGAAACCTCT 1547  
DB ATCTACCCAGAACCTTACATGAGCA---GAGATTATTTCCCAATCATGTAACCTCAT 1557  
QY 1548 CAGACGCCCATGCGCCCGAGATCTCATCTGCTCTCATCC 1586  
DB CGAAGCCCTGGAGCCCGCGCTTCATCAGCGCTCCACC 1596

## RESULT 8

DQ053108  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

DQ053108  
Pan troglodytes Hc13650 gene, VIRTUAL TRANSCRIPT, partial sequence.  
DQ053108  
DQ053108.1 GI:66899055  
GSS.  
Pan troglodytes (chimpanzee)  
Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Pan.  
1 (bases 1 to 996)  
Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,  
Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civeello,D.,  
White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.  
A Scan for Positively Selected Genes in the Genomes of Humans and  
Chimpanzees  
(ex) PLOS Biol. 3 (6), E170 (2005)  
15869325  
2 (bases 1 to 996)  
Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,  
Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civeello,D.,  
White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.  
Direct Submission  
Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.

FEATURES  
source Location/Qualifiers  
1..996  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..>996  
/locus\_tag="HC13650"

gene  
ORIGIN

Query Match 26.8%; Score 783; DB 11; Length 996;  
Best Local Similarity 78.9%; Pred. No. 1.2e-109;  
Matches 786; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY 450 GCCACTCCCAAGAGACGGGGTGGAACTGCTTGGCAGCCCGGTGGAGACACATCTTC 509  
DB 1 GCCACTCCCAAGAGACGGGGTGGAACTGCTTGGCAGCCCGGTGGAGACACATCTTC 60

QY 510 TCCTCCCAACACGCTAAATTCACGAGCGCATCGTAAGAGGAAGACGCTGGTGCCCC 569  
DB 61 TCCTCCCAACACGCTAAATTCACGAGCGCATCGTAAGAGGAAGACGCTGGTGCCCC 120

QY 570 AGAGATCAACATTTCTGTGATCAGAGTGAGGGTCCCTGCTCTCGATGACTTCTTGA 629  
DB 121 AGAGATCAACATTTCTGTGATCAGAGTGAGGGTCCCTGCTCTCGATGACTTCTTGA 180

QY 630 TACCCCTGATGACTGATTAACGTGGATGATCGAGACCCCGATGAGACCGACTC 589  
DB 181 TACCCCTGATGACTGATTAACGTGGATGATCGAGACCCCGATGAGACCGACTC 240

QY 690 GCTGGAGTTCTCTGGGAATGGCAACGAACTGGAGTGGGAAGACACACCCCGTGCCAC 749  
DB 241 GCTGGAGTTCTCTGGGAATGGCAACGAACTGGAGTGGGAAGACACACCCCGTGCCAC 300

QY 750 CGCAAGACATCCCGGGAGACGCGGATCTATTATTTGGGAGCGGACGAGAGACGG 809  
DB 301 CGCAAGACATCCCGGGAGACGCGGATCTATTATTTGGGAGCGGACGAGAGACGG 360

QY 810 CAGCGCGCAACGCGGCTGTGGCGGACAGTGATCATCGGGGACGAGACACCGTAT 869  
DB 361 CAGTGGCGGACGAGCGCTGTGGCGGACAGTGATCATTTGGGAGCAAGACGAGACCGTAT 420

QY 870 AGACCTGCATGATCGGCTTTACATGAAGTGGTCAACCGAGGGTACTACGGCGA 929  
DB 421 AGACCTGCATGATCGGCTTTACATGAAGTGGTCAACCGAGGGTACTACGGCGA 480

QY 930 AGGCTCAAGCCATCATCTCTTTCGAGCGCTCTCTTCCTTTCAGACAGCAGCTCCCGA 989  
DB 481 AGGCTCAAGCCATCATCTCTTTCGAGCGCTCTCTTCCTTTCAGACAGCAGCTCCCGA 540

QY 990 CTACCACTACATCATGAGAACCTTCTCTGTACGTTCATCAGCAGCTTAGCTCTGCT 1049  
DB 541 CTACCACTACATCATGAGAACCTTCTCTGTACGTTCATCAGCAGCTTAGCTCTGCT 600

QY 1050 GGCTGAGGACTACATGATCGTGTACCTGAAACGCTGCAACCGCCCGGAGGATGCTG 1109  
DB 601 GGCTGAGGACTACATGATCGTGTACCTGAAACGCTGCAACCGCCCGGAGGATGCTG 660

QY 1110 AATCGGCTGCTGAAGAGTGTCTACAGATGATCGACCGAGGTTCCGGAAGAACTGAA 1169  
DB 661 AATCGGCTGCTGAAGAGTGTCTACAGATGATCGACCGAGGTTCCGGAAGAACTGAA 720

QY 1170 GTCTTTGATCATCGTCCACCCCTCGTGTTCATTCCGACTGTGCTCTCTCGCC 1229  
DB 721 NNN 780

QY 1230 TTTTCATCAGCGTCAAGTTTCATCAACAAGATCCAGTACGTGACAGCTTGAAGACCTGGA 1289  
DB 781 NNN 840

QY 1290 GCAACTCATCCCTATGGAACACGCTCCAGATCCCAAGATCGGCTCTCGAATACGAAAGAGGA 1349

DB 841 NNN 1409  
QY 1350 AAGACTGAAGCGCAGGAGGAGCGCGAGGCGCCCGGAGTTTGTCTCCCGAGGTC 1409  
DB 901 AAGACTGAAGCGCAGGAGGAGCGCGAGGCGCCCGGAGTTTGTCTCCCGAGGTC 960  
QY 1410 TGAAGAGAGCCAGAGGTTGGCAACCACTGGAACACAG 1445  
DB 961 TGAAGAGAGCCAGAGGTTGGCAACCACTGGAACACAG 996

RESULT 9  
CN420201 748 bp mRNA linear EST 16-MAY-2004  
328775518 GRN\_ES Homo sapiens CDNA 5', mRNA sequence.  
CN420201  
CN420201.1 GI:47407795  
EST.  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 748)  
Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,  
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,  
Lebkowski, J. and Stanton, L.W.  
Transcriptome characterization elucidates signaling networks that  
control human ES cell growth and differentiation  
Nat. Biotechnol. 22 (6), 707-716 (2004)  
15146197  
Contact: Brandenberger R  
Regenerative Medicine  
Gen Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@gen.com  
Insert Length: 748 Std Error: 0.00.

FEATURES  
source Location/Qualifiers  
1..748  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, cell lines H1, H7, and H9"  
/clone\_lib="GRN ES"  
/note="oligo dt primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

ORIGIN  
Query Match 25.6%; Score 748; DB 7; Length 748;  
Best Local Similarity 100.0%; Pred. No. 2.9e-104;  
Matches 748; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 GCTACTCGAGGAGCG 96  
DB 1 GCTACTCGAGGAGCG 60

QY 97 CGCGCTGGGAGCG 156  
DB 61 CGCGCTGGGAGCG 120

QY 157 GGGCGCTGGCGACGCTCGCGCTGGGTGACCTTCTCGATGCAAGATCGCGCCCTCGGAGC 216  
DB 121 GGGCGCTGGCGACGCTCGCGCTGGGTGACCTTCTCGATGCAAGATCGCGCCCTCGGAGC 180

QY 217 ATCTCTTCTCTCTAGGCTCTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 276  
DB 181 ATCTCTTCTCTCTAGGCTCTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240

```
QY 277 GGCAGGGCTCGCCTTTGTTTGGCCAGTAAGGAGGAGGCTGTCTCAGCTGCAGAGGGTC 336
Db 241 GGCAGGGCTCGCCTTTGTTTGGCCAGTAAGGAGGAGGCTGTCTCAGCTGCAGAGGGTC 300
QY 337 ATCCCTGCTTCAAGCCAGTGCCTTCTCCAGCTCCCATGGGACCAACGAGGACGCTC 396
Db 301 ATCCCTGCTTCAAGCCAGTGCCTTCTCCAGCTCCCATGGGACCAACGAGGACGCTC 360
QY 397 CGGATGGAACCTGTGACGTGAAGGAGGAATGCGAGGACGAAGATCTTCCAGGCCACTC 456
Db 361 CGGATGGAACCTGTGACGTGAAGGAGGAATGCGAGGACGAAGATCTTCCAGGCCACTC 420
QY 457 CCAGAAGAGACGGGGTGGAACTGCTTGGCAGCCCGGTGGAGACATCTCTCTCC 516
Db 421 CCAGAAGAGACGGGGTGGAACTGCTTGGCAGCCCGGTGGAGACATCTCTCTCC 480
QY 517 AACACGCTAAATTTCAACGAGCGCATCTGAAGAGGAAGACGCTGTGGCCCGCAGATC 576
Db 481 AACACGCTAAATTTCAACGAGCGCATCTGAAGAGGAAGACGCTGTGGCCCGCAGATC 540
QY 577 AACATTTCTGATCAGAGTGAAGGGTCCCTGCTGTCTGATGACTTCTTGGATACCCCT 636
Db 541 AACATTTCTGATCAGAGTGAAGGGTCCCTGCTGTCTGATGACTTCTTGGATACCCCT 600
QY 637 GATGACCTGGATATTAACGTGGATGATCGAGACCCCGATGAGACCGACTCGCTGGAG 696
Db 601 GATGACCTGGATATTAACGTGGATGATCGAGACCCCGATGAGACCGACTCGCTGGAG 660
QY 697 TTCCTGGGGAATGGCAACCACTGGAGTGGGAAGACGACACCCCGTGGCCACCGCAAG 756
Db 661 TTCCTGGGGAATGGCAACCACTGGAGTGGGAAGACGACACCCCGTGGCCACCGCAAG 720
QY 757 AACATGCCCGGGACACGCGGATCTAT 784
Db 721 AACATGCCCGGGACACGCGGATCTAT 748
```

## RESULT 10

```
BG912726
LOCUS 602807930P1 NCI CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4940286
DEFINITION 5', mRNA sequence.
```

```
ACCESSION BG912726
```

```
VERSION BG912726.1 GI:142933202
```

```
KEYWORDS EST.
```

```
SOURCE Homo sapiens (human)
```

## ORGANISM

```
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
1 (bases 1 to 865)
```

```
NIH-MGC http://mgs.nci.nih.gov/
```

```
National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
Unpublished (1999)
```

```
Contact: Robert Strausberg, Ph.D.
```

```
Email: cgapbs-remail.nih.gov
```

```
Tissue Procurement: David N. Louis, M.D.
```

```
cDNA Library Preparation: Life Technologies, Inc.
```

```
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

```
DNA Sequencing by: Incyte Genomics, Inc.
```

```
Clone distribution: MGC clone distribution information can be
```

```
found through the I.M.A.G.E. Consortium/LLNL at:
```

```
http://image.llnl.gov
```

```
Plate: LLAM10878 row: k column: 07
```

```
High quality sequences stop: 821.
```

```
Location/Qualifiers
```

```
1..865
```

```
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
```

```
/clone="IMAGE:4940286"
```

```
/tissue_type="anaplastic oligodendroglioma with lp/19q
```

```
loss"
```

## FEATURES

```
source
```

## RESULT 11

```
CV024962
```

```
LOCUS
```

```
DEFINITION
```

```
2512 Full Length cDNA from the Mammalian Gene Collection Homo
```

```
CV024962
```

```
LOCUS
```

```
DEFINITION
```

```
/lab_host="DH10B (T1 phage-resistant)"
```

```
/clone_lib="NCI CGAP_Brn67"
```

```
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
```

```
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
```

```
Average insert size 2.3 kb. Constructed by Life
```

```
Technologies. Note: this is a NCI_CGAP Library."
```

## ORIGIN

```
Query Match 25.0%; Score 730; DB 2; Length 865;
Best Local Similarity 95.7%; Pred. No. 1.5e-101;
Matches 804; Conservative 0; Mismatches 30; Indels 6; Gaps 5;
```

```
QY 839 CAGTGATCATCGGGAGCAAGAGCACCGTATAGACCTGCACATGATCCCGCTTACATCA 898
Db 1 CAGTGATCATCGGGAGCAAGAGCACCGTATAGACCTGCACATGATCCCGCTTACATCA 60
QY 899 AAGTGGTCAACCCACGAGGGTACTACGGGAAAGGCTCAAACGCCATCATCGTCTTGGCAG 958
Db 61 AAGTGGTCAACCCACGAGGGTACTACGGGAAAGGCTCAAACGCCATCATCGTCTTGGCAG 120
QY 959 CCTGCTTCTTCCAGACAGCAGCCTCCCGACTACCACTACATCATGGAAGACCTCTTCC 1018
Db 121 CCTGCTTCTTCCAGACAGCAGCCTCCCGACTACCACTACATCATGGAAGACCTCTTCC 180
QY 1019 TGTACGTCAATCAGCAGCTTATAGAGCTCTGCTGGCTGAGGACTACATGATCGTGTACTCA 1078
Db 181 TGTACGTCAATCAGCAGCTTATAGAGCTCTGCTGGCTGAGGACTACATGATCGTGTACTCA 240
QY 1079 ACGGTGCCACGCCCCGGCGGAGGATGCTTGGAAATCGGCTGGTGAAGAGTCTACCCAGA 1138
Db 241 ACGGTGCCACGCCCCGGCGGAGGATGCTTGGAAATCGGCTGGTGAAGAGTCTACCCAGA 300
QY 1139 TGATCGACCGGAGGTTGGGAAACCTGAAAGCTCTTGCATCATCGTCCACCCCTGTGGT 1198
Db 301 TGATCGACCGGAGGTTGGGAAACCTGAAAGCTCTTGCATCATCGTCCACCCCTGTGGT 360
QY 1199 TCATTCGGAGCTGTGCTGGCCATCTCTCGCCCTTTTCATCAGGCTCAAGTTCATCAACAAGA 1258
Db 361 TCATTCGGAGCTGTGCTGGCCATCTCTCGCCCTTTTCATCAGGCTCAAGTTCATCAACAAGA 420
QY 1259 TCCAGTACGTGCACAGCTTGGAAAGACCTGGAGCAACTCATCCCTATGGAACACGTCACAGA 1318
Db 421 TCCAGTACGTGCACAGCTTGGAAAGACCTGGAGCAACTCATCCCTATGGAACACGTCACAGA 480
QY 1319 TCCAGACCTGCTCTGCAATACGAGAGGAAGACTGAAGCCAGGAGGAGAGCGGCA 1378
Db 481 TCCAGACCTGCTCTGCAATACGAGAGGAAGACTGAAGCCAGGAGGAGAGCGGCA 540
QY 1379 GGGCCCGAGCGGAGTTGTGTGCTGCCAGGTCTGAAGAGAAGCCAGAGTGGCACCAGTGG 1438
Db 541 GGGCCCGAGCGGAG-TTGTGTGACAGGTCTGAAGAGAGGAGGAGGAGGAGGAGTGG 599
QY 1439 AAAACAGGTCTGCTTGGTCTCAGAAAGATCAGGAAACAAAGCATGTCTGAGCGGACGTGA 1498
Db 600 AAAACAGGTCTGCTTGGTCTCAGAAAGATCAGGAAACAAAGCATGTCTGAGCGGACGTGA 659
QY 1499 GCATAACAAAGACATGGAAGAGATTCAGAA--TGCCAGAAAACCTCTGTTCAGACGCC 1556
Db 660 GCATAACAAAGACATGGAAGAGATTCAGAAAGATTCAGAAAGATTCAGAAAGATTCAGAAAG 719
QY 1557 ACTGGCCCCAGATCTCATCTCTGCTCATCTGAGTCCCAATCTTCCAGAGGAGGAGGCCGCC 1616
Db 720 ACTGG-ACCAGATCTCATCTCTGCTCAT-CTGAGTCCCATCTTTCAGAGGAGGAGGCCGCC-CC 776
QY 1617 CTCGGTTCATCTCTGAAACCCAGCATCTTTTCAGCTGCTTGAACAACTTGTATTTTTTT 1676
Db 777 CTCGGTGCATCTCTGAAACCCAGCATCTTTTCAGCTGCTTGAACAACTTGTATTTT 836
```

```
CV024962 699 bp mRNA linear EST 20-AUG-2004
```

```
LOCUS 2512 Full Length cDNA from the Mammalian Gene Collection Homo
```

```
DEFINITION
```

sapiens cDNA 5' similar to BC026217, mRNA sequence.

ACCESSION CV024962  
VERSION CV024962.1 GI:51482851  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 699)  
AUTHORS Rual, J.F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S.,  
Dricot, A., Li, N., Rosenberg, J., Lamesch, P., Vidalain, P.O.,  
Clingingsmith, T.R., Hartley, J.L., Esposito, D., Cheo, D., Moore, T.,  
Simmons, B., Sequerra, R., Bosak, S., Doucette-Stamm, L., Le Peuch, C.,  
Vandenhaute, J., Cusick, M.E., Albaladejo, J.S., Hill, D.E. and Vidal, M.  
TITLE Human ORFeome Version 1.1: a Platform for Reverse Proteomics  
JOURNAL Genome Res. (2004) In press  
COMMENT Contact: Vidal M  
Marc Vidal Laboratory  
Dana Farber Cancer Institute  
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA  
Tel: 617 632 5180  
Fax: 617 632 5739  
Email: Marc.Vidal@dfci.harvard.edu  
ORF Sequence tag (OST) of Gateway Entry construct. Each cloned ORF  
results from a PCR reaction using an MGC full-length cDNA as  
template DNA and ORF specific primers  
PCR Primers  
FORWARD: ATGGGGACCACCGAGCCAC  
BACKWARD: CAGGACATCGTGTCTCTGAT  
Insert length: 699 Std Error: 65.00  
Plate: 11058 row: 09 column: B  
Seq primer: ACTGGCGCTGTTTACAACTCGTGACTGGGAARAC  
High quality sequence start: 98  
High quality sequence stop: 698  
POLYA=No.

## FEATURES

source Location/Qualifiers  
1..699  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="mixed"  
/clone\_lib="Full Length cDNA from the Mammalian Gene  
Collection"  
/notes="Vector: mixed; The ORFs were PCR amplified from the  
MGC (Mammalian Gene Collection) as of April 2004 and  
cloned by recombinational Gateway cloning into pDONR223  
Donor vector. Reference : MGC (Mammalian Gene Collection)  
Program Team, Generation and Initial Analysis of more than  
15,000 Full-Length Human and Mouse cDNA Sequences. PNAS,  
2002, 99(26), 16899-16903"  
ORIGIN  
Query Match 23.9%; Score 696.4; DB 7; Length 699;  
Best Local Similarity 99.7%; Pred. No. 2.1e-96;  
Matches 697; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 373 ATGGGGACCACCGAGCCACGCTCCGGATGGAAACGCTGAGCTGAAGGAGGAATGGCAG 432  
DB 1 ATGGGGACCACCGAGCCACGCTCCGGATGGAAACGCTGAGCTGAAGGAGGAATGGCAG 60  
QY 433 GACGAAGATCTTCCAGGCCACTCTCCAGAGAGACGCGGGGTGGAACTGCTTGGCAGCCCG 492  
DB 61 GACGAAGATCTTCCAGGCCACTCTCCAGAGAGACGCGGGGTGGAACTGCTTGGCAGCCCG 120  
QY 493 GTGGAAGACATCTCTCTCTCCCAACACGCTAAATTTCAACGAGCGCATCGTAGAGG 552  
DB 121 GTGGAAGACATCTCTCTCTCTCCCAACACGCTAAATTTCAACGAGCGCATCGTAGAGG 180  
QY 553 AAGACGCTGTGGCCCCCAGAGATCAACATTTCTCTGATCAGATGAGGGGTCCTCGTGG 612  
DB 181 AAGACGCTGTGGCCCCCAGAGATCAACATTTCTCTGATCAGATGAGGGGTCCTCGTGG 240

QY 613 TCCGATGACTTCTTGGATACCCCTGATGACCTTGATATTAACGTGGATGACATCGAGACC 672  
DB 241 TCCGATGACTTCTTGGATACCCCTGATGACCTTGATATTAACGTGGATGACATCGAGACC 300  
QY 673 CCCGATGAGACCGACTCGCTGAGATTCTTGGGAATGGCAAGAACTGGAGTGGGAAGAC 732  
DB 301 CCCGATGAGACCGACTCGCTGAGATTCTTGGGAATGGCAAGAACTGGAGTGGGAAGAC 360  
QY 733 GACACCCCGTGGCCACCGCAAGAACATGCCCGGGGACAGCGCGGATCTATTTCGGGAC 792  
DB 361 GACACCCCGTGGCCACCGCAAGAACATGCCCGGGGACAGCGCGGATCTATTTCGGGAC 420  
QY 793 GGCACGCGGAGGACGCGGACGCGCCCAACGCGGCGCTGTGGCGGACAGTGTATTCGGG 852  
DB 421 GGCACGCGGAGGACGCGGACGCGCCCAACGCGGCGCTGTGGCGGACAGTGTATTCGGG 480  
QY 853 GAGCAAGAGCACCGTATAGACCTGCACATGATCCGGCCCTTACATGAAGTGTGTACCCAC 912  
DB 481 GAGCAAGAGCACCGTATAGACCTGCACATGATCCGGCCCTTACATGAAGTGTGTACCCAC 540  
QY 913 GGAGGGTACTACGCGGAAGGCTCAACGCCATCATCTTTCGACGCTTCCCTTCCA 972  
DB 541 GGAGGGTACTACGCGGAAGGCTCAACGCCATCATCTTTCGACGCTTCCCTTCCA 600  
QY 973 GACAGACGCTTCCCGACTTACCACATCATGAGGAACCTCTTCTCTGTACGTCTATCAGC 1032  
DB 601 GACAGACGCTTCCCGACTTACCACATCATGAGGAACCTCTTCTCTGTACGTCTATCAGC 660  
QY 1033 AGCTTAGAGCTCTGCTGGCTGAGGACTACATGATCGTG 1071  
DB 661 AGCTTAGAGCTCTGCTGGCTGAGGACTACATGATCGTG 699

## RESULT 12

LOCUS DR002011  
DEFINITION TC120419 Human fetal brain, large insert, PCMV expression library  
Homo sapiens cDNA clone TC120419 5' similar to Homo sapiens  
KIAA1872 protein (K1AA1872), mRNA sequence.  
ACCESSION DR002011  
VERSION DR002011.1 GI:66261884  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 714)  
AUTHORS Birckett, C., Cho, J., Gau, Y., Hamer, R., Kelly, S., Kovacs, K., Liu, L.,  
Liu, X., Porter, J., Sachs, A., Shu, Y., Sun, Z., Wong, J., Wu, M.,  
Zhang, X., Jay, G. and He, W.  
High-throughput cloning of full-length human cDNAs directly from  
cDNA libraries optimized for large and rare transcripts  
Unpublished (2005)  
CONTACT: Kovacs, KF  
High Throughput cDNA Cloning  
Origene Technologies, Inc. (www.origene.com)  
6 Taft Court, Suite 100, Rockville, MD 20850, USA  
Tel: 301 340 3188  
Fax: 301 340 8606  
Email: cDNA@origene.com

This EST submission is part of an on-going human full-length  
cloning project at Origene Technologies, Inc.  
Please contact Origene for access.  
Origene Technologies, Inc.  
6 Taft Ct. Suite 100  
Rockville, MD 20850  
Tel: (301) 340-3188  
http://www.origene.com

Seq primer: PCMV6 5prime forward vector primer, Origene  
Technologies Inc.  
Location/Qualifiers  
1..714  
source





```
Db 421 TAACTCAGATGTCAAGCCACCGGCAAAACCCCGTCAATACCTCCCAAGGACTGAGAT 480
QY 2267 ATGTGGACCTCACTGTCTCCCAACCCAGCGTCAGGCTGGGACACGCCCAACGCTGTTCG 2326
Db 481 ATGTGGACCTCACTGTCTCCCAACCCAGCGTCAGGCTGGGACACGCCCAACGCTGTTCG 540
QY 2327 GGTGTGAA-CAGCAGAGGCTCAGAAACTGG-CTCTGAAATAGCAGACCTTAGCAAGAGGA 2384
Db 541 GGTGTGAA-CAGCAGAGGCTCAGAAACTGG-CTCTGAAATAGCAGACCTTAGCAAGAGGA 600
QY 2385 AGATACAGGTTATCGGGCGTTTTCAGTGTTCAGAGTCA-TTCGGGAAGATAAATCCAGT 2443
Db 601 CGATACAGGTTATCGGGCG-TTCAGTGTTCAGAGTCACTTCGGGACCATCAATCCAGT 659
QY 2444 GCCTCGCCGCGCAGCCACCTGTCATTTCAAGCTTCGACGAGCGGTTCTTGTTCGGGAGGCA 2503
Db 660 GCGCTGGCGGTAGCCACTGCATTTCAAGCTTCGACGAGCGGTTCTTGTTCGGGAGGCA 717
QY 2504 AATTTCCTAGGAAAAGAGACAGACTTTTCTTAATGTGTGTTCGTCGAAATGCGGATCACTGGT 2563
Db 718 AATTTCCTCGGCCACACAGACCGACTTTCTCT-ATGTGTCTCAAAATGCGGATCAC-GGT 775
QY 2564 CAGATGACCTCTAGAGCACTGAGCTCCCTGTCTCTGGA 2602
Db 776 CAGATGACCTCCAGAGC-CTGAGCTCCCTGTCTCTGGA 813
```

```
RESULT 14
AL138250 650 bp mRNA linear EST 04-SEP-2003
LOCUS DKFZP547E187.r1 547 (synonym: hfbri) Homo sapiens cDNA clone
DEFINITION DKFZP547E187 5', mRNA sequence.
ACCESSION AL138250
VERSION AL138250.1 GI:6854931
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
TITLE Hominidae; Homo.
JOURNAL 1 (bases 1 to 650)
COMMENT Ansoorge,W., Wirkner,U., Mewes,W., Weil,B. and Wiemann,S.
EST (Ansoorge,W., Wirkner,U., Mewes,H.W., Weil,B. and Wiemann,S.)
Unpublished (1999)
Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No 5' sequence available.
This clone (DKFZP547E187) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
```

```
FEATURES
source
1..650
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZP547E187"
/tissue_type="brain"
/dev_stage="fetal"
/lab_host="Xl-2blue"
/clone_lib="547 (synonym: hfbri)"
/notes="Vector: pAMP1; Site_1: NotI; Site_2: SalI"
```

## ORIGIN

```
Query Match 22.2%; Score 648.4; DB 1; Length 650;
Best Local Similarity 99.8%; Pred. No. 4.5e-89;
```

```
Matches 649; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1846 ACAAGATTCAAGACCCCTTCTCTTGTGTGTACCCGCTCAGGCTTGAGGACACAGACACCC 1905
Db 1 ACAAGATTCAAGACCCCTTCTCTTGTGTGTACCCGCTCAGGCTTGAGGACACAGACACCC 60
QY 1906 ACCGCAACCCCGGCTGCGGTCTGCTCTTCTGTGCTCTTCTCCAGAAATGCGGCTTC 1965
Db 61 ACCGCAACCCCGGCTGCGGTCTGCTCTTCTGTGCTCTTCTCCAGAAATGCGGCTTC 120
QY 1966 AGACCTAGAGCTCAACCCCTTATGAGGCGCCACCTCTCTGGGGTAGCTCTCTGACCTCCA 2025
Db 121 AGACCTAGAGCTCAACCCCTTATGAGGCGCCACCTCTCTGGGGTAGCTCTCTGACCTCCA 180
QY 2026 CCTTATGTCCAAATTTCAACCCATATGTTTTCATTTGACCCGCCCCCTTCTCGCTCATA 2085
Db 181 CCTTATGTCCAAATTTCAACCCATATGTTTTCATTTGACCCGCCCCCTTCTCGCTCATA 240
QY 2086 ATGACACCCAGCTCTCTTTGAGAGGATCAGAGCCCATTTGCACAAGAGAGCCGCTGCCAAC 2145
Db 241 ATGACACCCAGCTCTCTTTGAGAGGATCAGAGCCCATTTGCACAAGAGAGCCGCTGCCAAC 300
QY 2146 CATCTTGTCTCCGATTTGCAAAATGACACCCAGTAATCTAGAAATCTCAAGCCCT 2205
Db 301 CATCTTGTCTCCGATTTGCAAAATGACACCCAGTAATCTAGAAATCTCAAGCCCT 360
QY 2206 TTAACCTCAGATGTCAGGCGGCAAAACCCGCTCAATACCTCCACCAAGGAATGAGA 2265
Db 361 TTAACCTCAGATGTCAGGCGGCAAAACCCGCTCAATACCTCCACCAAGGAATGAGA 420
QY 2266 TATGTGGACCTCACTGCTCTCCCAACCCAGCGTCAGGCTGGGACACGCCAGCTGTTC 2325
Db 421 TATGTGGACCTCACTGCTCTCCCAACCCAGCGTCAGGCTGGGACATGCCAACGCTGTTC 480
QY 2326 GGGTTGGAACACAGAGGCTCAGAACTGGCTCTGAAATAGGACAGACCTAGCAAGGAA 2385
Db 481 GGGTTGGAACACAGAGGCTCAGAACTGGCTCTGAAATAGGACAGACCTAGCAAGGAA 540
QY 2386 GATACAGGATATCGGCGGTTTCAGTGTTCACAAAGTCATTCGGGAAGATAAATCCAGTGC 2445
Db 541 GATACAGGATATCGGCGGTTTCAGTGTTCAGAACTATTCGGGAAGATAAATCCAGTGC 600
QY 2446 GCTGCCCGCAGCACCTGCTCAATTCAAAGCTTGGACAGCGGGTCTTGTTC 2495
Db 601 GCTGCCCGCAGCACCTGCTCAATTCAAAGCTTGGACAGCGGGTCTTGTTC 650
```

```
RESULT 15
CN420203 639 bp mRNA linear EST 16-MAY-2004
LOCUS 1700600176974 GRN_PRENBU Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN420203
ACCESSION CN420203
VERSION CN420203.1 GI:47407797
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
TITLE Hominidae; Homo.
JOURNAL 1 (bases 1 to 639)
COMMENT Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J. and Stanton,L.W
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
15146197
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
```

Email: rbrandenberger@eron.com  
Insert Length: 639 Std Error: 0.00.  
Location/Qualifiers  
1. .639

FEATURES  
source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cell, retinoic acid and  
mitogen-treated hES cell line H7"  
/clone\_lib="GRN PRENEU"  
/notes="oligo dt primed, full-length enriched cDNA library  
from hES cell line H7 (p29) maintained in feeder-free  
conditions. Embryoid bodies were generated in the presence  
of all-trans retinoic and mitogens."

## ORIGIN

Query Match 21.9%; Score 639; DB 7; Length 639;  
Best Local Similarity 100.0%; Pred. No. 1.2e-87;  
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 52 GGGGCGCGCAGCGCGCGCGATGAGAGGGGGCGCAGCCGCGCGCTGGGGAGCCC 111  
Db 1 GGGGCGCGCAGCGCGCGCGATGAGAGGGGGCGCAGCCGCGCGCTGGGGAGCCC 60  
QY 112 ACCGCTAACCTGACACCCACCCACCCCTGTCACAAAAGAGCTGGGGGCGCTGGCCACGT 171  
Db 61 ACCGCTAACCTGACACCCACCCACCCCTGTCACAAAAGAGCTGGGGGCGCTGGCCACGT 120  
QY 172 CGCCCTGGGTGACCTTCTCGGATGAGATCGCCCTCGGAGCATCTCTTCCTCCTA 231  
Db 121 CGCCCTGGGTGACCTTCTCGGATGAGATCGCCCTCGGAGCATCTCTTCCTCCTA 180  
QY 232 GGCTCTGAAGCGCGCGGAGCGTGCAGCGATGCCCGAGCTGACCCGCGCAGGGCTCGCCTT 291  
Db 181 GGCTCTGAAGCGCGCGGAGCGTGCAGCGATGCCCGAGCTGACCCGCGCAGGGCTCGCCTT 240  
QY 292 TGTTCGCGATGAAGAGGAGAGCGTGTCTCAGCTGCAGAGGGGTCAATCCCTGCTTCAAGC 351  
Db 241 TGTTCGCGATGAAGAGGAGAGCGTGTCTCAGCTGCAGAGGGGTCAATCCCTGCTTCAAGC 300  
QY 352 CAGTGCTCTTCCAGCTCCATGGGGACCAACGAGCCACGCTCCGGATGGAAAACGTG 411  
Db 301 CAGTGCTCTTCCAGCTCCATGGGGACCAACGAGCCACGCTCCGGATGGAAAACGTG 360  
QY 412 GAGCTGAAGGAGGAATGGCAGGACGAAGATCTTCCAGGCCACTCCCGAAGAGAGCGGG 471  
Db 361 GAGCTGAAGGAGGAATGGCAGGACGAAGATCTTCCAGGCCACTCCCGAAGAGAGCGGG 420  
QY 472 GTGGAACTGTTGGCAGCCCGGTGGGAAGACACATCTCTCTCCCAACGCTAAATTTTC 531  
Db 421 GTGGAACTGTTGGCAGCCCGGTGGGAAGACACATCTCTCTCCCAACGCTAAATTTTC 480  
QY 532 AACGGAGCGCATCGTAAGGAGAGAGCGTGGTGGCCCGCAGAGATCAACATTTCTTGGAT 591  
Db 481 AACGGAGCGCATCGTAAGGAGAGAGCGTGGTGGCCCGCAGAGATCAACATTTCTTGGAT 540  
QY 592 CAGAGTGAGGGGTCCCTGCTGCTCCGATGACTTCTTGGATACCCCTGATGACCTGGATAT 651  
Db 541 CAGAGTGAGGGGTCCCTGCTGCTCCGATGACTTCTTGGATACCCCTGATGACCTGGATAT 600  
QY 652 AACGTGGATGACATCGAGAGCCCCCGATGAGACCGACTCG 690  
Db 601 AACGTGGATGACATCGAGAGCCCCCGATGAGACCGACTCG 639

Search completed: March 5, 2006, 06:52:44  
Job time : 11158 secs

**THIS PAGE BLANK (USPTO)**